THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Extremely Strong Contiguous Hydrogen Bonding Arrays

by

Patrick I T Thomson

Degree of Doctor of Philosophy

The University of Edinburgh

May 2013

Dedicated to Prof. Hamish McNab

11/3/49 - 15/11/10

Contents

Abstract

When multiple hydrogen bonds lie in-plane and parallel to each other in close proximity, they experience additional positive or negative secondary electrostatic interactions. When a pair of molecules are arranged such that every hydrogen bond acceptor is on one molecule and every hydrogen bond donor is on another, the positive secondary electrostatic interactions are maximised, and thus the association constant of the complex is enhanced.

This thesis will present the development of a family of quadruple hydrogen bonded complexes containing only positive secondary interactions, which confers unprecedented stability. The complexes are sufficiently stable to maintain strong binding in polar solvents such as acetonitrile and can be switched "on" and "off" with acid and base. They will be developed into synthons for acid-base responsive supramolecular recognition, for use in stimuli-responsive supramolecular polymers and gelators.

Declaration

I declare that this thesis is my own composition, that the work described herein has been carried out by me unless otherwise stated, and that it has not been submitted to any previous application for a higher degree.

This thesis describes the results of research carried out between the 1st August 2009 and the 1st July 2012. The work was carried out in the School of Chemistry, the University of Edinburgh, under the supervision of Prof. Hamish McNab and Prof. David Leigh.

Signed: ……

Date: …….

Meetings Attended and Posters and Talks given

22nd IUPAC Conference on Physical Organic Chemistry, Durham, UK, 9th-13th September 2012. Selected to give 2.5 minute "flash" talk as well as present poster.

University of Edinburgh PhD Prize Symposium (29th June 2012, Edinburgh). Gave presentation; won prize for best speaker.

7th Annual International Symposium for the Advancement of the Chemical Sciences (12th - 15th June 2012, Edinburgh). Stewarding with full attendance.

University of Edinburgh Organic Division Seminar (2nd-4th April 2012, Loch Tay). Gave presentation; won prize for best speaker.

RSC Chemical Nanoscience Spring Symposium (14th March 2012, Newcastle). Presented poster.

7th Annual International Symposium for Macrocyclic and Supramolecular Chemistry (29th Jan - 2nd Feb 2012, Dunedin, New Zealand). Presented poster.

RSC Physical Organic Chemistry Postgraduate Meeting (14th September 2011, Macclesfield). Presented poster; won prize for best poster.

Syngenta Agrochemicals Training Day (8 April 2011, Glasgow)

RSC Regional Organic Meeting (15th December 2010, Edinburgh). Presented poster.

Attended eight other regional organic chemistry meetings (2009-2012)

Acknowledgements

I would like to thank my supervisor, Prof. David Leigh for his support and wisdom, and for the opportunity to work with him for the past three years.

I would not be where I am today without the late Prof. Hamish McNab, joint supervisor of this work until his death at the end of 2010. His door was always open to everyone, even when time was short. He listened to, advised, guided, and cared about everyone who worked for him, and through knowing him we were all inspired. He believed that we had the capacity to succeed, even when we didn't, and we are all poorer for his passing.

I am also indebted to Dr. Alison Hulme, Dr. Scott Cockroft, and Dr. Peter Kirsop, who were there with support and guidance throughout the process of personal and professional development these last three years. I also wish to thank Dr. Perdita Barran and Prof. Mark Bradley for offers of office space while writing this thesis, and Lilian Monahan for being a friend and ally.

I am grateful to the technical staff of the University of Edinburgh for all the help and advice, particularly Juraj and Lorna for NMR, Alan for mass spectroscopy, Phil, Ron and the Admiral for technical and building services, Stuart for glassblowing, Davey and Stuart for mechanical engineering, Donald for electronics, and last but not least Tim, John, Raymond and Derek in stores.

Recognition to everyone involved in undergraduate teaching and demonstrating, I salute you for working full-time on an endeavour that exhausted me after six hours a week: Peter, Steve, Rehana, Donald, Patrick, Elaine and Kirsty.

For McNab Group members past: Andreas, Craig, Emma, Lore, Martha, Mary, Ran, Tom, and Will.

For Leigh Group members past and present: Adam, Armando, Barney, Barry, Bartek, Bea, Chris, Clint, Craig, Daniel, Dave H, Guillaume, Gus, Jeff, Jenny, Jon, Jon, Louise, Malcolm, Marcus, Max, Miriam, Paul, Sonja, Steffen, Steven, Stuart, Ula and Victor.

For the friends who kept me sane while writing this: Iain, Shanana, Will, Pete, Emma, Luke, Sarah, Kirsty, Sarah-Jane, Meg, Benji, Sarah-Lee, Tim, Emma, Andrew, Dani, Nicola, Jack, Alex, Michael, Stuart, Lynn, Kelly, Tim, Alister. Clergy and members of P's & G's: Dave, Rich, Vanessa, Dean, Emma, Fi, John, Marc, Mark, Helen, Carrie, Mary, Tim, Mary, Katie, Chris, Lizzie, Liz, Jamie, Jonno, Cathy, Cassie, Hazel. Other good friends too many to list here. Particular shout-out to Mike and Lindsay; friends for many years before this work started, and for many to come.

For the people I met or got to know after submitting this thesis, who got me through my corrections: Lisa, Colin, Kate, Jing, Anna-Maria, Lauren, Feny, Ben, Richard, Charlotte, Damion, Andrew, Ilka.

For the celebrities, visionaries and authors I did not know personally, but who produced things I read or watched between working and sleeping: Lauren Faust, Warren Spector, John Carmack, Pendleton Ward, Matthew, Mark, Luke, John, Ronald D Moore, J Michael Straczynski, Peter Molyneux, Kan Gao, Rob Bell, Darren Korb, Gabe Newell, Oliver Sacks, John Lennon, Notch, C S Lewis, Primo Levi, Marcus Borg, Tom Wright, Stephen Hillenburg, Matt Groening, Tim Burton, Paul, John Cleese, J R R Tolkien, George Lucas, Walt Disney.

Mum.

List of Abbreviations

A	hydrogen bond acceptor
APCI	Atmospheric Pressure Chemical Ionisation
aq	aqueous
BArF	tetrakis(3,5-bis(trifluoromethyl)phenyl)borate
BuLi	butyllithium
C	Celcius
calcd	calculated
δ	chemical shift
D	hydrogen bond donor
DCM	dichloromethane
DME	1,2-dimethoxyethane
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
eq	equivalents
ESI	Electrospray Ionisation
Et_2O	diethyl ether
EtOH	ethanol
g	grams
h	hours
HBA	hydrogen bond acceptor
HBD	hydrogen bond donor
IPA	isopropyl alcohol

LDA	Lithium diisopropylamide
M	Molar
m/z	mass/charge ratio
MHz	Megahertz
MeCN	Acetonitrile
MeOH	Methanol
min	minutes
mL	milliliters
mp	melting point
MS	Mass Spectroscopy
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMR	Nuclear Magnetic Resonance
ppm	parts per million
quat	quaternary
rt	room temperature
Tf	triflate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TLC	Thin Layer Chromatography
UPy	2-ureido-4[1*H*]-pyrimidinone
UV/vis	Ultraviolet/visible

Standard abbreviations for NMR peak multiplicities were used as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br.s = broad singlet, dd = doublet of doublets, td = triplet of doublets, pn = pentet, dt = doublet of triplets, br = broad.

Layout of the Thesis

Chapter I outlines the existing research into linear hydrogen bonding arrays of triple, quadruple, and higher orders, and the importance of secondary electrostatic interactions and preorganisation to the strength of their complexes. Chapters II and III describe the development and characterisation of various strong hydrogen bonding arrays with no negative secondary interactions. They are each presented in the form of articles that have recently been published or prepared for publication in peer-reviewed journals, and have been reformatted to ensure consistent presentation. These chapters are also reproduced in Appendix I in their published formats where available. Chapter IV describes work to incorporate strong hydrogen bonding into a supramolecular polymer.

Chapter I

Introduction to Hydrogen Bonding Arrays

Acknowledgements

Drs Barry Blight and Craig Robertson are very gratefully acknowledged for proof-reading this chapter.

1.1 Introduction

1.1.1 The hydrogen bond

In the field of molecular recognition and self-assembly, one of the most widely-used non-covalent interactions is the hydrogen bond. Known for over 100 years,[1,2] the hydrogen bond is ubiquitous in nature as the enforcer of DNA base pairing[3-5] and in the secondary structure of proteins.[6] Hydrogen bonds are usually of moderate strength, being weaker than covalent bonding but stronger than dispersion forces. The definition of a hydrogen bond is wide-ranging,[7] but the vast majority of cases are covered by the 2011 IUPAC recommended definition[8] of a hydrogen bond:

The hydrogen bond is an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation.

The current work is concerned with hydrogen bonding between pairs of molecules, where the molecule or molecular fragment X–H is referred to as the Hydrogen Bond Donor (HBD) and the molecule or molecular fragment to which the HBD forms a bond is referred to as the Hydrogen Bond Acceptor (HBA).

1.1.2 The binding constant

Because hydrogen bonds are relatively weak, the binding energy is usually not quantified directly in the same way as a covalent bond. Because there is normally no major kinetic barrier to hydrogen bond formation, they can be studied in equilibrium between different species in the same way as some other weak intermolecular interactions. Consider two species, A and B, which form a complex, C. At equilibrium, the reaction scheme is:

$$A + B \rightleftharpoons C$$

And the association constant (K_a) is given by the ratio of A, B, and C as follows:

$$K_a = \frac{[C]}{[A][B]}$$

Therefore, having some way of observing the position of this equilibrium can give values of K_a. For example, if the NMR chemical shift of a proton changed during an equilibrium titration, then it could be tracked, plotted, and data fitted to give an association constant. The Gibbs free energy of the A-B interaction can then be calculated as follows:

$$\Delta G = -RT \ln K_a$$

Sometimes it is most helpful to discuss hydrogen bonding in terms of association constants and sometimes by free energy of binding, and the two are generally used interchangeably.

1.1.3 Hydrogen bond arrays

When two hydrogen bonds lie in-plane and parallel to each other in close proximity* (separated by not more than one carbon atom, or about 2 – 3 Å), significant additional energetic factors come into play. When adjacent hydrogen bonds are oriented in the same direction, they experience an additional attractive interaction. Conversely, adjacent hydrogen bonds that are oriented in the opposite direction experience a repulsive interaction.

These additional interactions are termed secondary interactions, to distinguish them from the primary interaction of the hydrogen bond itself. As an illustration of the importance of secondary interactions to the overall complex strength, consider the interaction between 2-aminoadenine and thymine, **1•2** (Figure 1.1, left). There are 3 primary hydrogen bonds, in an "ADA" arrangement and with a total energy of ≈ -11.5 kJ mol^{-1}.

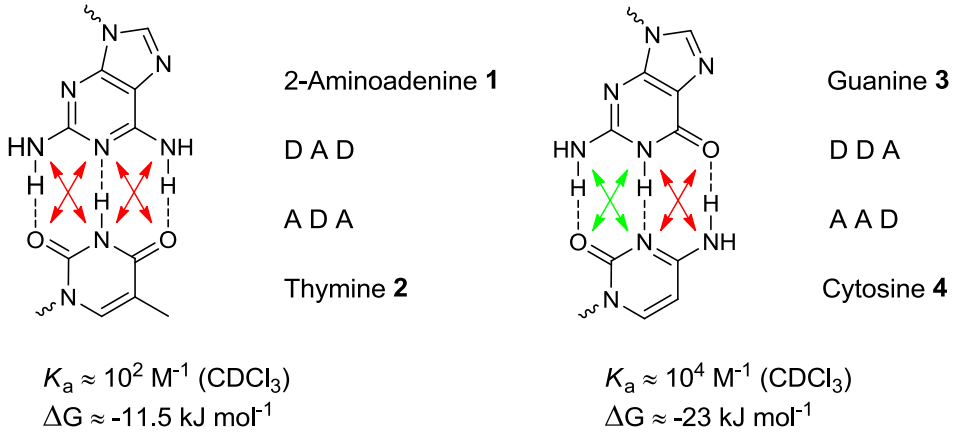


Figure 1.1 | Secondary interactions in natural and unnatural base pairing.[13,14] Positive secondary interactions are shown in green and negative secondary interactions in red. "A" is used to denote the Acceptor end of the hydrogen bond, and "D" for the Donor.

* Widely-spaced[9,10] and helical[11] arrays are important to the understanding of hydrogen bonding but do not usually offer insight into secondary electrostatic interactions and will not be discussed here. Anslyn has recently reported a helical array with a significant energetic contribution from secondary electrostatic interactions.[12]

However, when contrasted with guanine-cytosine base pair **3•4** from DNA (Figure 1.1, right), there are some discrepancies that cannot be explained by considering the number and type of hydrogen bonds. In both cases, there are three hydrogen bonds, and in both cases they have similar types of atoms participating in the hydrogen bond itself. The only difference is that one of the hydrogen bonds has been "swapped over", exchanging the positions of the HBA and HBD, changing the overall arrangement from an ADA•DAD to a DDA•AAD. As a consequence, two of the four secondary interactions have switched from negative to positive, and this single change is responsible for a doubling of the interaction energy in non-polar solvents from -11.5 to -23 kJ mol^{-1}. From this simple discrepancy, it is evident that secondary interactions need to be taken into account in order to even qualitatively predict the strength of contiguous hydrogen bonding arrays.

The importance of secondary interactions has been appreciated since 1990, when Jorgensen first described the effect in his Secondary Interaction Hypothesis.[15,16] Along with a more comprehensive investigation by Schneider,[17] the Secondary Interaction Hypothesis describes in detail the effect of secondary interactions on overall complex strength. The authors proposed that the stability of a hydrogen bonded complex depends on two factors: The number of hydrogen bonds, responsible for *ca.* 8 kJ mol^{-1} of stabilisation each, and interactions between neighbouring hydrogen bonds (secondary electrostatic interactions) responsible for *ca.* ± 2.9 kJ mol^{-1} per interaction. However, Jorgensen's model was empirically derived from a small number of examples and using it quantitatively has been cautioned against in several computational studies from Popelier[18] and Zimmerman.[19]

Even if no precise predictions can be made without computational methods, it is apparent that secondary interactions have a vital role in determining the strength of a hydrogen bonding complex. We will therefore present a review of some recent developments in complementary hydrogen bonding motifs, specifically with a view to illustrating the importance of secondary electrostatic interactions. For this reason, we will only consider complexes with three or more primary hydrogen bonds.

1.2 Triple hydrogen bond arrays

The most famous example of a complex held together by three hydrogen bonds is also the oldest; the guanine-cytosine base pair of DNA and RNA is held together by an AAD•DDA arrangement of hydrogen bonds and has been used by nature for billions of years. The complementary arrangement of base pairs was elucidated by Crick, Watson and Franklin in 1953 shortly after proposing the structure of the double helix.[20] Triple hydrogen bonded complexes have subsequently been studied in various recognition patterns,[13,14] with the first comprehensive work coming from Zimmerman in 1995 where he presented examples of triple hydrogen bonded arrays in each possible permutation (That is, ADA•DAD, AAD•DDA, and AAA•DDD).[21]

Certainly, triple hydrogen bonded systems have been much easier to synthesise than quadruple hydrogen bonded complexes and we will not attempt to provide a comprehensive review, but instead look at selected examples, specifically those with no positive secondary interactions (ADA•DAD) contrasted with those with all positive secondary interactions (AAA•DDD). The secondary interaction hypothesis predicts that these will have very different binding constants, with ADA•DAD at around 10^2 M^{-1} and AAA•DDD being some three orders of magnitude stronger at 10^5 M^{-1}.

1.2.1 Triple arrays with no positive secondary interactions

Hydrogen bonding pairs that carry all negative secondary interactions (that is, ADA•DAD) have been known since the inception of the field, appearing in early work by Jorgensen[15,16] and Zimmerman.[22] Of particular interest is Zimmerman's ADA motif **6**, which templates a variety of DAD motifs with modest binding. The structure differs only slightly from an AAA•DDD complex (Section 1.2.2, Figure 1.5) but is at least four orders of magnitude weaker.

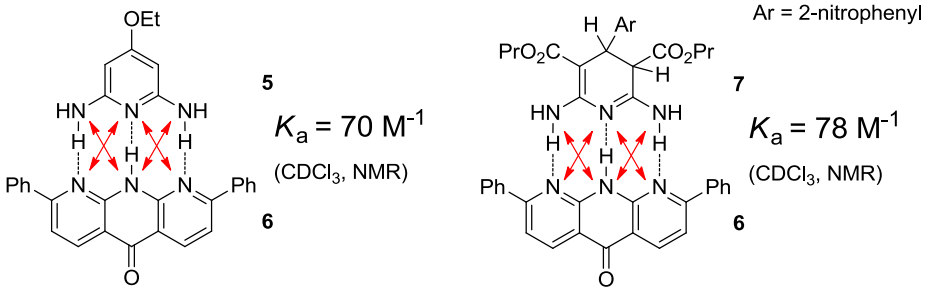


Figure 1.2 | ADA motif based on an oxidised phenanthridine.

Binding constants for ADA•DAD complexes are known to range from a K_a of 65-900 M^{-1}, spanning an entire order of magnitude of binding constants but only 6 kJ mol^{-1} when considered energetically. Zimmerman makes the case that this is actually a relatively narrow range of binding constants, since minor variations in experimental conditions can account for slight differences in the energy.[23] The most widely-used DAD motif comes from acetamide-substituted pyridines. Utilised by Lehn in his "rigid rods" (Chapter IV, Figure 4.2),[24] these motifs have been elaborated and investigated by other researchers.

Wilson developed[25] a conformation-independent complementary DAD based on a urea-pyrimidine motif as an offshoot from work on AAD motifs. In this case, the donor's rigidity derived from an internal hydrogen bond that locked the ring structure in the correct confirmation (Figure 1.3, H_a). The conformational independence arises because either pyrimidine HBA can participate in the internal hydrogen bond, but this does not alter the arrangement of the recognition motif. In this array the principle of

preorganisation is powerfully illustrated; not only is the arrangement conformation-independent, but it is actually locked into place by an internal hydrogen bond, reducing the energetic penalties to association. Although only binding with a modest K_a of 56 M^{-1}, the authors attribute this to competition from self-association.

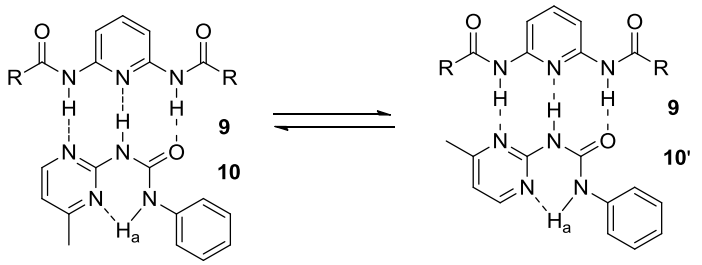


Figure 1.3 | 10 and 10' both present an ADA arrangement.

Ośmiałowski has carried out experimental and theoretical studies[26-28] on acylaminopyridine donors, varying the steric bulk of acyl substituents to give binding constants ranging from 27-380 M^{-1}. Even a change in conformation of the steric bulk altered the binding mode (Figure 1.4).

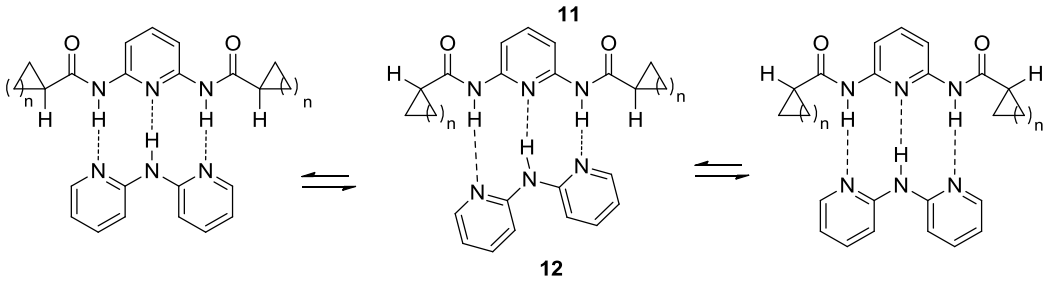


Figure 1.4 | Rotational equilibrium in bulky acyl groups push an ADA further away from the binding face of the acylpyridine DAD

Generally, triple hydrogen bonded complexes with no positive secondary interactions are of limited utility due to their relatively weak binding, usually being reported and evaluated as undesirable isomers or tautomers of other, stronger systems.

1.2.2 Triple arrays with only positive secondary interactions

The first reported example of a triple hydrogen bonded complex with only positive secondary interactions came in 1992 from Zimmerman and Murray,[22] using a previously-known[29,30] 1,9,10-anthyridine as a triple HBA. When evaluating complex **13•14** (Figure 1.5), they noted that the association constant was beyond the limits of detection by NMR at $K_a \geq 10^5$ M^{-1}. In contrast, the closely-related ADA-DAD complex **7•6** has four repulsive secondary interactions and an association constant of just 78 M^{-1}. This striking difference in strength serves to illustrate the importance of secondary interactions to the overall complex strength.

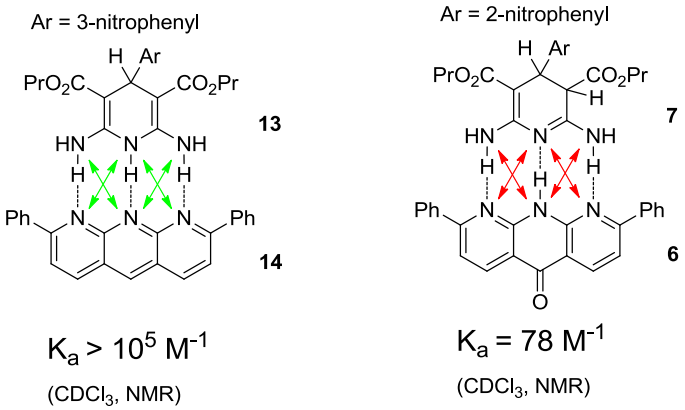


Figure 1.5 | The first reported AAA-DDD, and the analogous ADA-DAD illustrating the tremendous influence that secondary interactions can have on binding strength.

Note that **7** exists entirely as the 3,4-dihydro form shown, even when not templated by **6**. However, **13** exists as a 67:33 equilibrium between the shown 1,4-dihydro and the undesired 3,4-dihydro forms in the absence of **14**. In the presence of **14**, however, the structure of **13** was as drawn here, indicating that the AAA•DDD interaction thus created was sufficient to drive the tautomeric equilibrium entirely to one side.

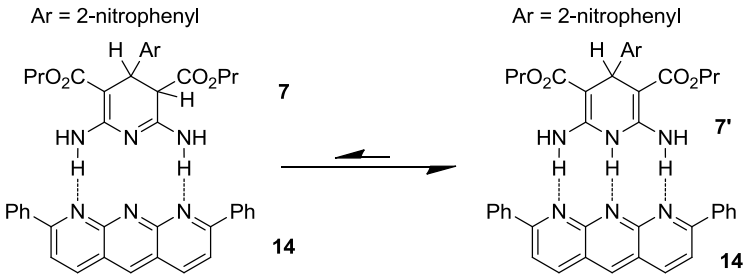


Figure 1.6 | Tautomerisation preference is reversed in the presence of a H-bonding template.

Interestingly, Zimmerman also reported[22] that **7** completely converts to the 1,4-dihydro isomer in the presence of the AAA partner **14** as shown (Figure 1.6).

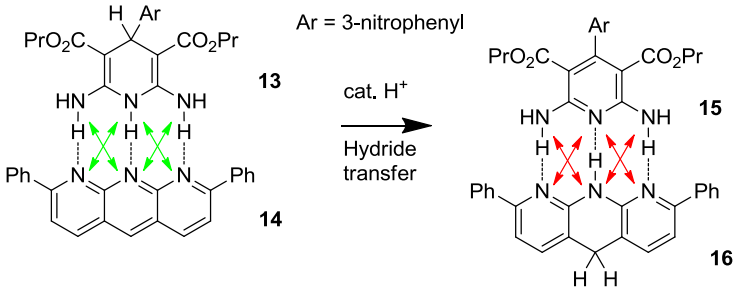


Figure 1.7 | Anthyridine instability in the presence of trace amounts of acid.

Compound **14** also suffered from significant stability problems, undergoing hydride transfer to give **16** in the presence of catalytic traces of acid (Figure 1.7), and oxidation to **6** under ambient conditions.[31] Zimmerman later reported[21] a more stable derivative **17** with a 5-aryl substituent that halted oxidation (Figure 1.8), but was still prone to hydride transfer.

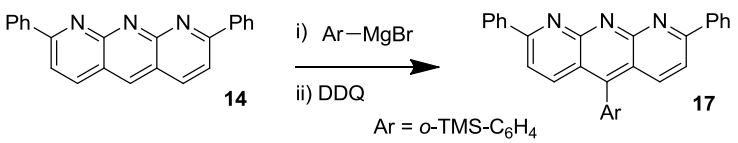


Figure 1.8 | Synthesis of a 5-substituted anthyridine.

At the same time as Zimmerman's initial work on the neutral complexes, Anslyn reported[32] the first cationic triple HBD, based on a protonated diaminonicotinate **18**. The cation was paired with the tetrakis[(3,5-trifluoromethyl)phenyl] borate (BArF) anion. This anion is much more solubilising than tetraphenylborate; **18** was soluble in $CDCl_3$ at concentrations of 1 mM, suitable for NMR studies. The anion is also extremely weakly interacting; a successive dilution experiment of **18** revealed a K_a of 37 M^{-1}, meaning that at the typical concentration used in a binding experiment, 0% of the cation was bound to the anion.

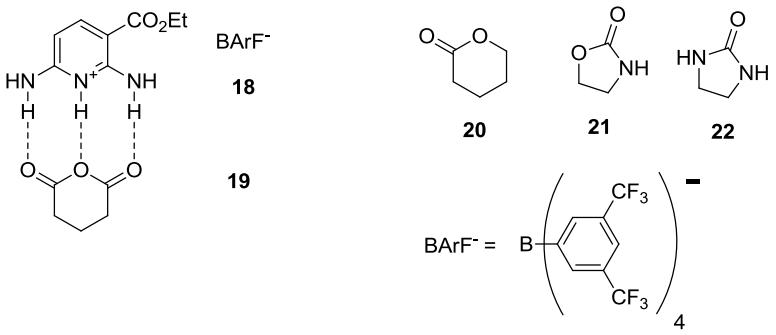


Figure 1.9 | Potentially, the first AAA•DDD array.

Anslyn was primarily interested in using **18** to coordinate to mono- and bidentate carbonyl compounds, but found that it complexed the cyclic anhydride **19** with a modest binding constant of $K_a = 6 \times 10^3$ M^{-1}. However, this donor also complexes urea **22** with a binding constant of 5×10^5 M^{-1}, nearly two orders of magnitude stronger. This urea is only capable of forming at most two bent hydrogen bonds, but is approximately 10 log units more basic than compound **19**. A trend including ester **20** and urethane **21** indicated that the interaction strength was independent of the number of hydrogen bonds or dipole moment, and most strongly correlated with basicity. In particular, **18•20** (DDD•AA) was over 10 times stronger than **18•19** making a specific AAA•DDD binding mode unlikely (Figure 1.9, left).

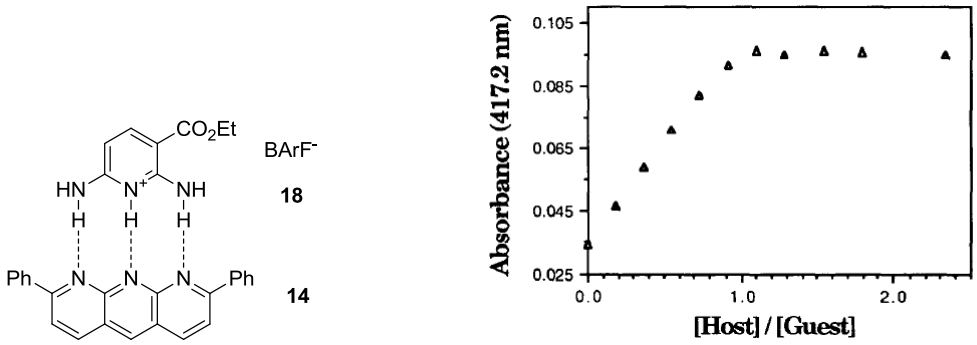


Figure 1.10 | Anslyn's cationic AAA-DDD and its binding isotherm (UV, 5.9 × 10^{-6} M, DCM).

Ansyln later[33] attempted to measure the binding strength of **18** with the heteroaromatic acceptor **14**, and found the complex to bind extremely strongly, being indicative of a true AAA•DDD mode (Figure 1.10, left) unlike **18•19**. Although evaluation was somewhat complicated by the presence of multiple equilibria, a sharply changing isotherm was observed (Figure 1.10, right) at 5.9 × 10^{-6} M indicating K_a > 1 × 10^6 M^{-1}. Although fluorescence spectroscopy can potentially probe lower concentration ranges and thus evaluate higher binding constants, Anslyn also reported that excited state proton transfer was occurring when measurement of the binding constant was attempted by fluorescence, and so no upper bound could be reported.

The first precisely-evaluated AAA•DDD came in 2007, when Leigh developed[34] a new triple HBA based on a strategy of tandem Suzuki coupling and imine condensation to generate fused pyridines.

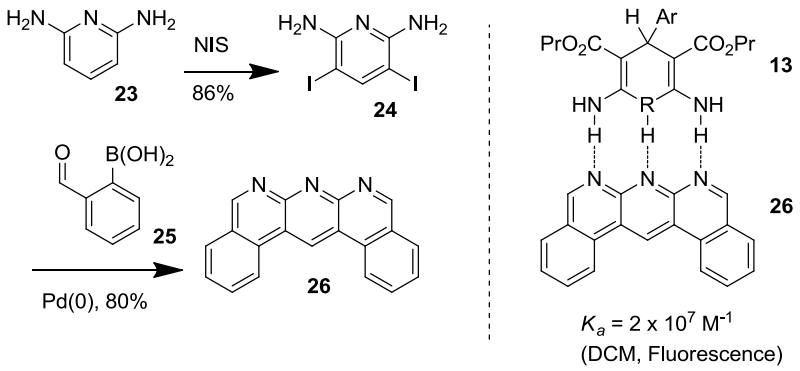


Figure 1.11 | Leigh's labile synthesis of a stable triple HBA.

Notably, the acceptor could be synthesised in two steps from commercial material in 70% overall yield, and was resistant to photo-degradation, photo-oxidation, and excited-state proton transfer. Use of fluorescence spectroscopy permitted the direct evaluation of the binding constant of 2×10^7 M^{-1} in DCM for **26•13**. They were also able to synthesise a structural analogue of **26** with only two HBAs and four rings (Figure 1.12, **27**). This can be directly compared to 1,8-naphthyridine **28**,[35] and the additional two rings increase the binding strength by a factor of 25. This is especially remarkable because **27** has a helical twist of 30° in the solid state, induced by steric crowding between protons in the cavity at the rear of the molecule.

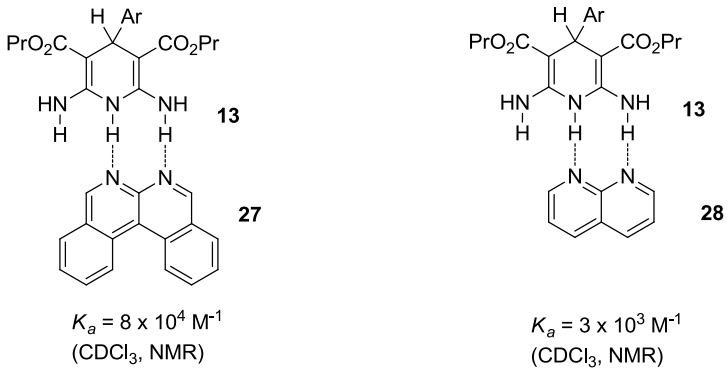


Figure 1.12 | Bonding in an AA•DDD array with different acceptor motifs.

Compound **26** was extremely fluorescent, allowing it to be detected at nanomolar concentrations and allowed the binding with cationic donors to be evaluated. The interaction of **26** with the 2,6-diaminopyridinium cation **29** was measured to be 3×10^{10} M^{-1}, which was at the time the highest reported binding constant for a small molecule hydrogen-bonded complex.[36]

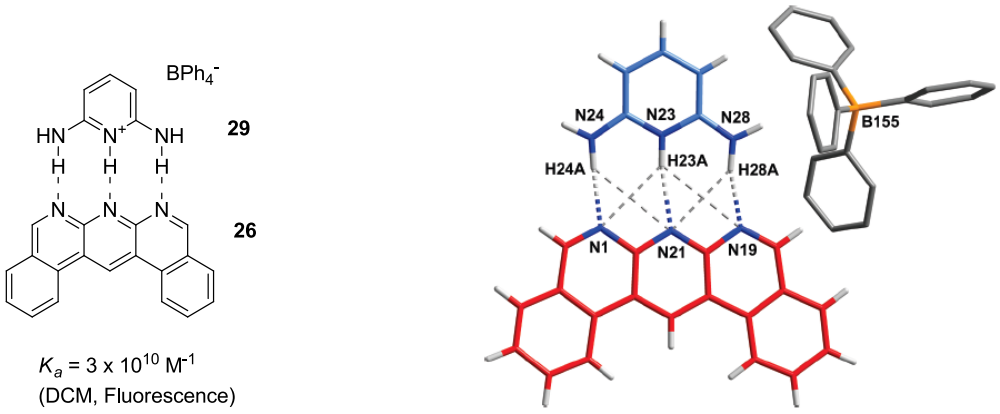
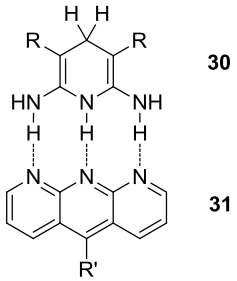


Figure 1.13 | The strongest known triple hydrogen bond complex, with crystal structure (BPh_4^- anion, solvent omitted for clarity).

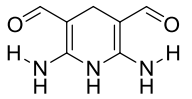
A crystal structure of the complex was obtained showing a slightly staggered conformation, which indicated that the hydrogen bonds did not need to be exactly linear, although this may be an artefact of the crystal packing and not representative of the solution-phase structure.

Factors other than primary and secondary interactions can contribute to the strength of hydrogen bond complexes, and some work has been done in this area by Boyd,[37] who investigated the strength of triple hydrogen bonded complexes between dihydropyridine **30** and phenanthridine **31** when various electron-donating and withdrawing substituents were incorporated into the donor molecule.



30

31



32

Figure 1.14 | Computational study of dihydropyridines as 3-HBDs.

For R = R' = H, they reported K_a = 5.5 × 10^5 M^{-1}, and found that non-conjugated electron-donating or withdrawing substituents on **30** had only a modest effect on K_a. Conjugated electron-withdrawing substituents capable of forming intramolecular hydrogen bonds (such as formyl, in **32**) were found to increase the planarity of the dihydropyridine ring and preorganise the amino groups with a stabilising effect of up to 40 kJ mol^{-1}, or 6 orders of magnitude (K_a up to 10^{11} M^{-1}). Several factors could not be anticipated in this computational study such as the already-established chemical instability of the chosen substrates, however it has again illustrated the importance of preorganisation.

1.3 Quadruple hydrogen bond arrays

Quadruple hydrogen bond arrays are a more recent development, but have been more widely applied than triple hydrogen bond arrays due to synthetic accessibility and the potential for greater energetic values. Sijbesma and Meijer in a 2003 review[38] calculated the expected values of the binding constants for each of the six possible arrangements of four hydrogen bonds.

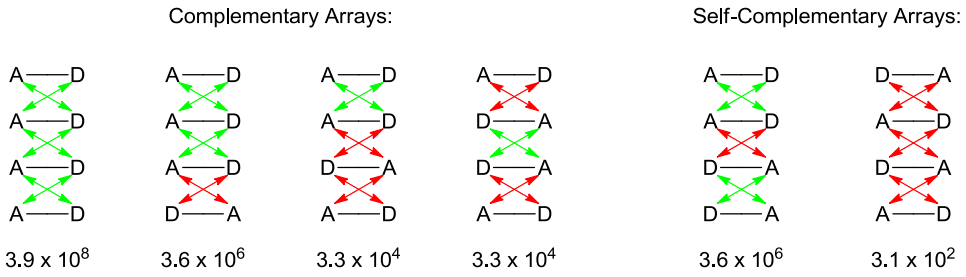


Figure 1.15 | Computed values of K_a (M^{-1}) for each quadruple hydrogen bonded array.[38]

It is interesting to note that the model calculates identical energies for some pairs of arrays: ADDA•DAAD and AADA•DDAD both have an expected association constant of 3.3×10^4 M^{-1} from two positive and four negative secondary interactions, whereas AAAD•DDDA and AADD•DDAA both have an expected association constant of 3.6×10^6 M^{-1} from four positive and two negative secondary interactions. In practice, these arrays usually have differing electronic and steric factors, hence the predictions of Figure 1.14 can be used only as a guide to expected trends.

The arrays are almost always heteroaromatic in nature, and Zimmerman has presented excellent reasoning for this in an older review[23] that covers in more detail the reasons for using heteroaromatic molecules; the feature that concerns us most is the capacity to form intramolecular hydrogen bonds for preorganisation.

1.3.1 Quadruple arrays with no positive secondary interactions

When the hydrogen bonding pattern of an array alternates (ADAD), all of the secondary interactions are negative and these complexes are usually very weak. In general, molecules that present an ADAD arrangement of hydrogen bonds will self-associate, however with rather low binding constants. Consequently, these compounds have not often been specifically presented in the current literature, as they are usually present as undesired tautomers of other, stronger arrangements.[23,38,39] There has been one interesting report[40] of systems specifically designed to bear an ADAD motif with no tautomerisation, and these systems furnished modest binding constants.

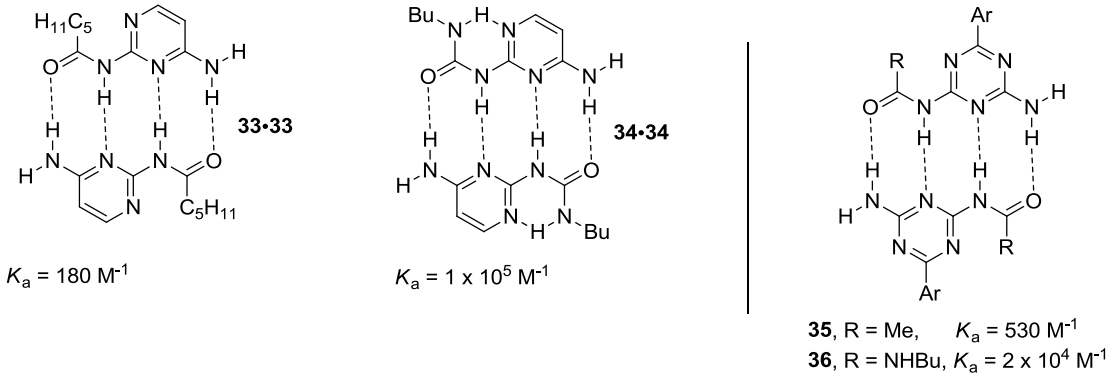


Figure 1.16 | Meijer's triazine and pyrimidine system with and without preorganisation mediated by internal hydrogen bonding. All K_a values measured by NMR in $CDCl_3$ at 298 K.

However weakly the compounds self-associate, a vital concept was demonstrated: preorganisation dramatically increases the binding constant. When acyl-amines are replaced with ureas, the presence of an additional intramolecular hydrogen bond dramatically increases the binding constant by several orders of magnitude, simply by locking the conformation of the carbonyl acceptor (Figure 1.15, centre).

1.3.2 Quadruple arrays with two positive secondary interactions

There are two possible permutations of quadruple hydrogen bonding with two positive and four negative secondary interactions, ADDA•DAAD and AADA•DDAD. These have both been realised, with the former being more well-explored.

The first report of an ADDA•DAAD comes from Zimmerman[41] in 1998; while developing the AADD motif **37** (Section 1.3.3, Figure 1.23), it was found to have some tautomers that presented an ADDA motif, and could be induced to exist solely in this form when paired with a diaminonaphthyridine-based DAAD **38**. In each case, the system is preorganised by an internal hydrogen bond as shown, which contributes significantly to the binding constant.

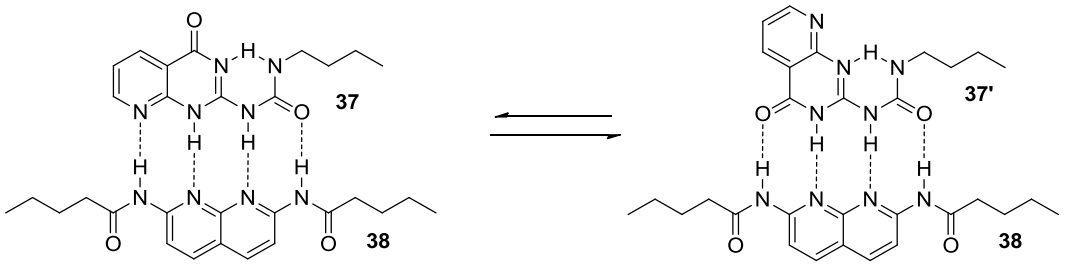


Figure 1.17 | Tautomers of Zimmerman's ADDA bonding with a diaminonaphthyridine DAAD

Although the strength of these complexes was not directly evaluated in non-polar solvents, competition experiments and experiments in DMSO/$CDCl_3$ mixtures indicated that this complex was stronger than predicted from the secondary interactions ($K_a \approx 1 \times 10^7$ M^{-1}). The authors proposed that the high strength was due to greater contributions from unusually strong primary interactions in **37•38**, a hypothesis supported by computations. Specifically, each "NH" donor is flanked by, or conjugated to, strongly electron-withdrawing groups.

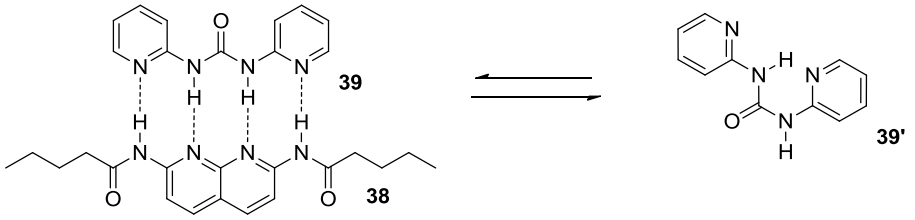


Figure 1.18 | Tautomers of di(2-pyridyl)urea, hindering binding as an ADDA

Other binding partners have been used with **38**; one such is di-pyridylurea **39**, which bound with an unexpectedly low K_a of 2200 M^{-1} in $CDCl_3$.[42] This was thought to be due to self-association and folding of the urea, highlighting again the importance of preorganisation to complex strength (Figure 1.18). When an appropriate binding partner is chosen, naphthyridine **38** is a versatile building block and has been tethered and incorporated into polymer blends.[43-45] It has also been used as a drop-in replacement for DNA base pairs and significantly increases the stability of short lengths of artifical DNA,[46] and it can also be derivatised in a way that includes the ability to switch complexation electrochemically (Chapter IV, Figure 4.7).[47]

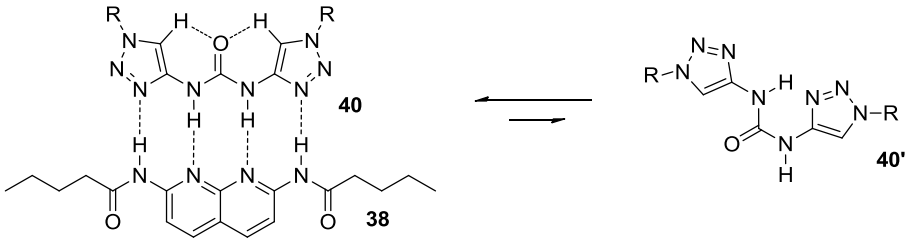


Figure 1.19 | Hisamatsu's triazole-based binding partner can still fold, but this is unfavourable and it cannot tautomerise.

Hisamatsu attempted to address tautomerisation with the development[48] of a triazole-based ADDA motif **40** that associates strongly with **38** ($K_a = 2.6 \times 10^5$ M^{-1} in $CDCl_3$). Although **40** can in principle also hydrogen bond intramolecularly, this is much less favourable than for **39**. Intriguingly, although unreported by Hisamatsu, the triazole protons appear to be contributing to preorganisation by acting as hydrogen bond donors towards the central carbonyl. Triazoles are known to act as hydrogen bond donors for anion recognition[49] and their use in hydrogen bonding arrays is covered in Chapter III of the current work.

A small amount of work has been carried out on the alternative arrangement of hydrogen bonds in this category, ADAA•DADD. In each case, acylaminopyridines were used as DADD motifs, which are capable of intramolecular folding to a greater or lesser extent. Several ADAA motifs have been developed based on naphthyridines[50,51] and triazine N-oxides[52] by Lüning and Sijbesma respectively. However, in each case the binding constants were modest, being 100 – 800 M^{-1} in $CDCl_3$ for **42•43** depending on the substituents, attributed to folding of the acylaminopyridine motif **43**. For the N-oxide **41**, no binding at all could be detected, which was taken to imply that the oxygen of an N-oxide is fundamentally unsuitable for hydrogen bonding applications.

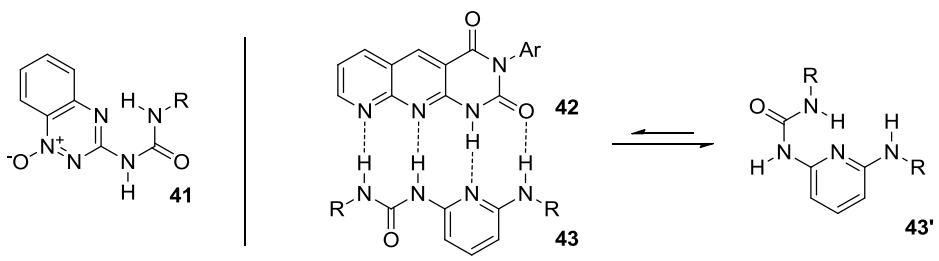


Figure 1.20 | A variety of AADA motifs, which experience modest binding to acylpyridine DDAD due to competition from intramolecular hydrogen bonding

1.3.3 Quadruple arrays with four positive secondary interactions

There are two possible permutations of quadruple hydrogen bonding with four positive and two negative secondary interactions, AADD•DDAA and AAAD•DDDA. The latter has not been explored, since the DDDA motif remains unknown. An AAAD motif has been synthesised but only ever used as an AAD motif in a triple hydrogen bond array.[21,53] The two known AAAD motifs had ether or ester oxygen acceptors. (**44** and **45**, Figure 1.21).

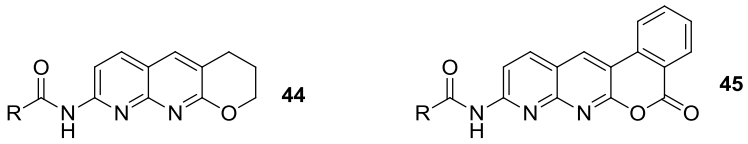


Figure 1.21 | Examples of an AAAD motif.

The remaining permutation AADD is self-complementary, and in this class of hydrogen bonding arrays the definitive examples are those based on 2-ureido-4[1*H*]-pyrimidinone **47** (colloquially known as UPy) developed by Sijbesma and Meijer. First reported in 1997,[54] the family has an exceptionally short synthesis, with most variants being assembled in a single step from commercial materials (Figure 1.22, left). Like Meijer's previous system (Section 1.3.1, Figure 1.16), the binding motif in UPy is preorganised with an intramolecular hydrogen bond.

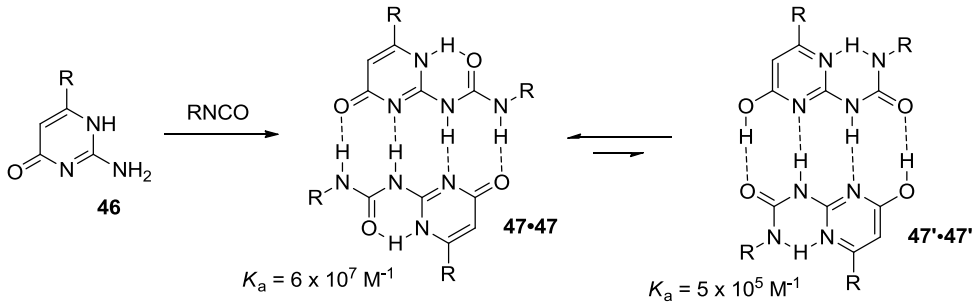


Figure 1.22 | Synthesis and subsequent self-association of UPy, in various tautomeric forms

One drawback of UPy is tautomerisation to a self-complementary ADAD motif, which has weaker binding (Figure 1.22, right). However, the self-complementary binding arrangement is preserved, and the system can tautomerise without breaking the arrangement of hydrogen bonds or grossly altering the geometry of any supramolecular systems held together with UPy linkages.

UPy's short synthesis, versatility, and at-the-time unprecedented binding constant of 6.0×10^7 M^{-1} in $CDCl_3$ meant that applications were rapidly developed, with uses in supramolecular polymers,[40,55], self-healing gels,[56] photoswitching polymers,[57] as a general-purpose synthon for hydrogen bonding,[58] dendrimer assembly,[59] self-assembly of superhydrophobic surfaces[60] and many others. Uses of UPy are numerous and exceed the scope of this introduction, but excellent reviews[61-67] and theoretical studies[68] exist. This benchmark material is likely the most widely-used small molecule hydrogen bonding recognition motif outside of DNA itself.

Attempting to address the issue of tautomerisation in heteroaromatic hydrogen bonding motifs, Zimmerman reported[41] an AADD which has strong self-association ($K_a > 10^7$ M^{-1}) regardless of the tautomeric form. However, drawbacks included a longer synthesis and an ADAD arrangement which was present as a *ca.* 5% component, complicating characterisation and impairing binding (Figure 1.23, right).

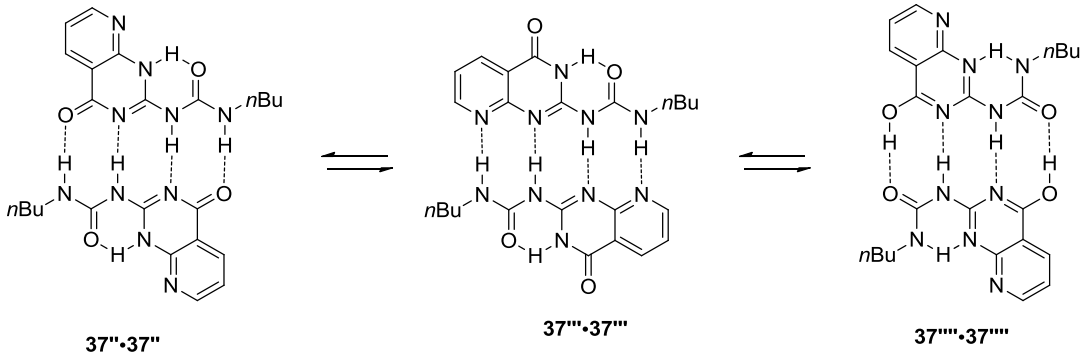


Figure 1.23 | Self-association of Zimmerman's AADD motif (centre), with the homodimer of one possible tautomer shown (left) and also an ADAD homodimer (right).

Another AADD motif, this time without tautomers, is the cytosine dimer developed by Hailes,[69] with a binding constant of $\geq 9 \times 10^6$ M^{-1}. Alkylation of the one of the

nitrogens of **48** with an R'' group prevented tautomerisation, and foldamers could be controlled by the choice of the R' group. When the group was a proton, an intramolecular hydrogen bond (heteroaromatic CH to carbonyl O) templated the quadruple hydrogen bonded arrangement **48''•48''**. However, when R' was replaced with a sterically undemanding fluorine atom, folding was promoted and the array was entirely in the form **48'•48'**.[70]

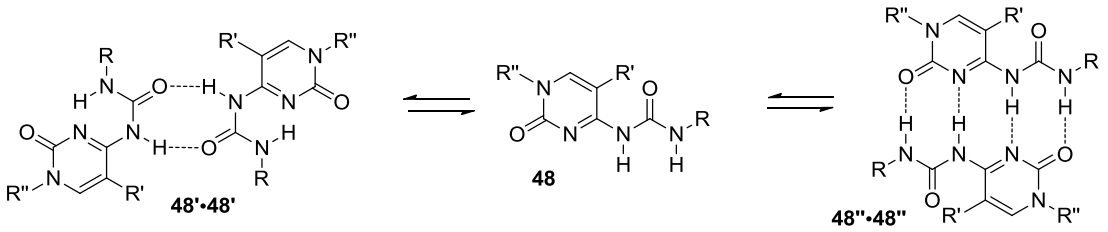


Figure 1.24 | Self-association of Hailes's cytosine dimer, in folded (left) and unfolded (right) states

Hisamatsu later developed[71] an AADD motif **49** that self-associates without folding or forming tautomers at all. The motif is based on a fused heterocyclic core with a urea substituent, crucially using a 5-membered ring to present a hydrogen bond acceptor site and simultaneously relieve steric bulk on the carbonyl of the urea in a similar manner as his triazole-based ADDA motif (Section 1.3.2, Figure 1.19)

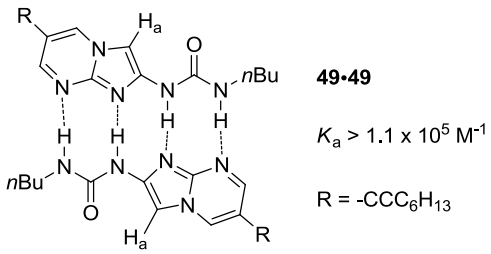


Figure 1.25 | Hisamatsu's non-tautomeric AADD motif

It could imagined that the one of the heteroaromatic C-H groups (H_a, Figure 1.25) has some degree of acidity and hydrogen bonds in a weak intramolecular fashion to the urea carbonyl, giving the complex some additional rigidity in preorganisation. The

complex is strong, but was only evaluated to the limit of NMR detection (K_a > 1.1 × 10^5 M^{-1}) making comparison with UPy difficult.

Finally, a quadruple array that definitively incorporates C-H bonds as hydrogen bond donors was developed by Yuan in 2011.[72] Compound **50** has an AA and DD motif, connected by a phenylene spacer that participates in hydrogen bonding as evidenced by the solid-state structure (Figure 1.26). The binding constant in chloroform is again reported at the NMR limit of K_a > 1 × 10^5 M^{-1}, but experiments carried out in polar solvent mixtures indicated comparable strength to UPy.

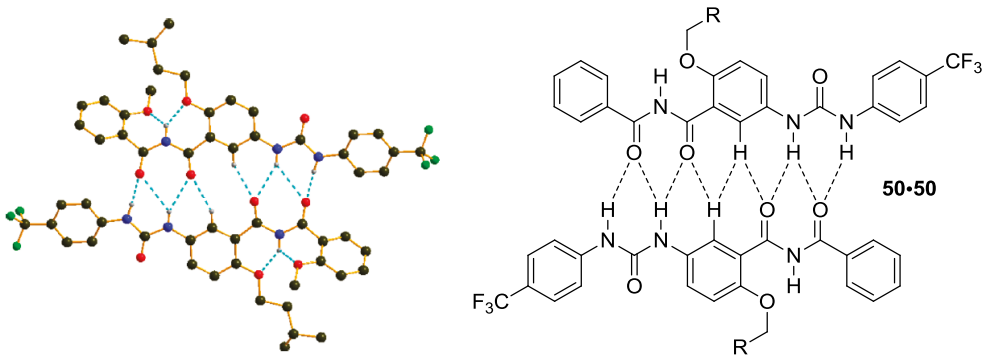


Figure 1.26 | A self-complementary AADD array with an aryl ring as a hydrogen bond donor.

1.3.4 Quadruple arrays with all positive secondary interactions

The highest predicted value of free energy of binding for a quadruple hydrogen bond complex is that for the AAAA•DDDD arrangement, containing six positive secondary interactions and no negative secondary interactions. In 2009, Lüning[73] published what appeared to be the first example of this, with sulfurane **50** acting as an AAAA, partnering with the cationic DDDD **51**.

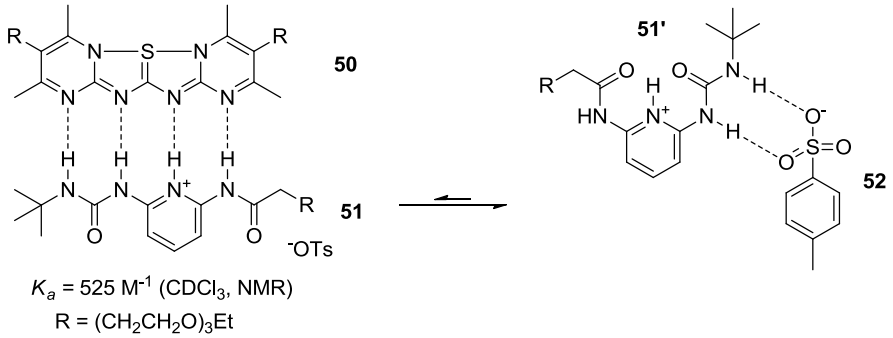


Figure 1.27 | The first AAAA•DDDD quadruple hydrogen bonded complex?

A very low association constant was reported for **50•51** (525 M^{-1}). Several explanations were offered for this poor binding constant, such as steric repulsion or proton transfer, but additional factors are likely to be in play, such as competition with the bidentate anion **52** and/or folded conformations as shown (Figure 1.27, right) leading to only non-specific ion-dipole interactions between **51** and **52**.

In 2011, Leigh published[74] a cationic AAAA•DDDD complex with $K_a > 3 \times 10^{12}$ M^{-1}. Chapter II of the current work pertains to the development and characterisation of this complex and it will not be discussed here.

1.4 Hydrogen bonded arrays with more than four hydrogen bonds

There are few examples of linear hydrogen bonding arrays with more than four contiguous hydrogen bonds; there have been some widely-spaced arrays using 6 hydrogen bonds, but these lack some or all secondary interactions, as well as suffering from a lack of conformational rigidity and subsequent folding.[9,10] There is one example of a contiguous array of six hydrogen bonds, from Corbin and Zimmerman in 2001.[42] They developed a modular approach to substituted naphthyridines that could be applied to synthesise **53** and **54**, which form six hydrogen bonds to each other. Multiple foldamers, undesirable self-association of both **53** and **54**, and a 1:2 ratio of negative to positive secondary interactions limited the strength of the array to 5×10^5 M^{-1} in $CDCl_3$.

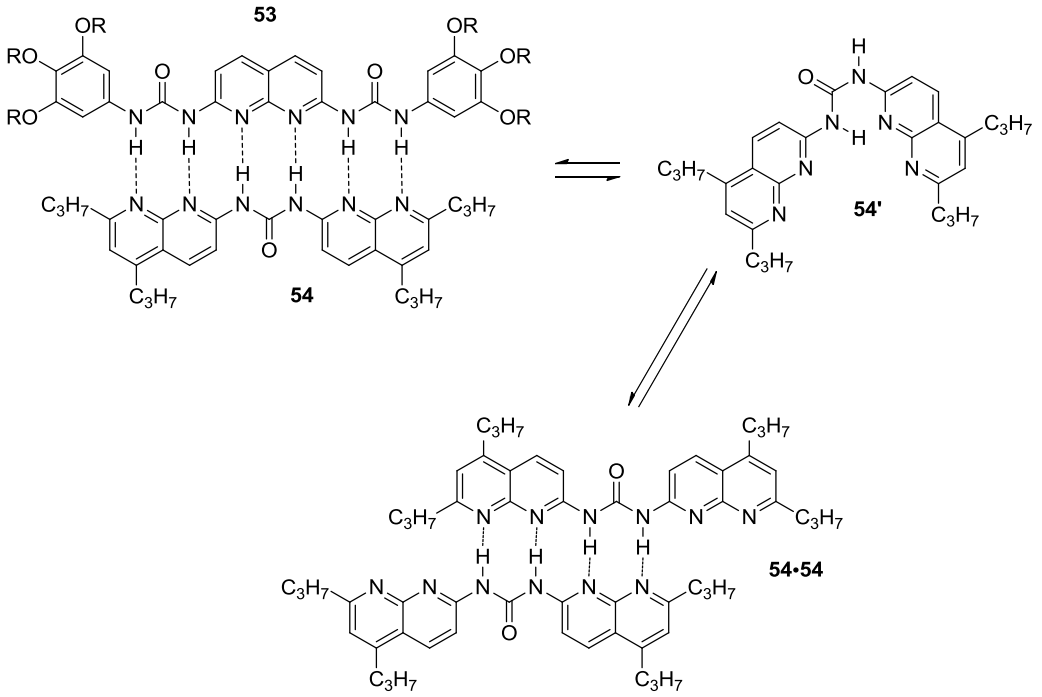


Figure 1.28 | A hexuple hydrogen bonded complex can undergo folding and self-association. Only one self-associated arrangement is shown, but many are possible.

1.5 Conclusion

Pairs of molecules with complementary hydrogen bonding motifs are able to interact favourably with each other; such patterns have been used since the advent of life itself to bind together DNA heteroduplexes. The particular patterns of hydrogen bonding used affect the free energy of interaction, and can generally be empirically predicted by considering the effects of secondary interactions, tautomerisation, and conformation. The trends presented in the secondary interaction hypothesis have been largely validated by experimental data, and cases where the free energy of interaction is anomalously low can usually be explained.

Understanding tautomerisation is particularly vital to predicting the interactions of hydrogen bonding heterocycles, since different tautomers can access different patterns of hydrogen bonding and different conformations. Intramolecular hydrogen bonds are particularly important for preorganisation, ensuring that hydrogen bond complexes are already in an optimal conformation for molecular recognition. Where no such preorganisation is present, the free energy of interaction with other hydrogen bonding motifs is usually low due to the necessity to overcome entropic and enthalpic factors that normally induce undesirable conformations.

1.5.1 Scope of the thesis

The present thesis explores the use of secondary electrostatic attractions and preorganisation to effect high binding constants in hydrogen bonding arrays. Chapter two is concerned with the development of a quadruple hydrogen bonding array of unprecedented stability. Chapter three focuses on modification of the array to incorporate heteroaromatic carbon-hydrogen bonds as hydrogen bond donors, and also the reversible stimuli-responsive decomplexation of hydrogen bond arrays. Finally, chapter four details work carried out to incorporate strong, stimuli-responsive hydrogen bonding into a supramolecular polymer.

1.6 References

(1) Moore, T. S.; Winmill, T. F. *J. Chem. Soc., Trans.* **1912**, *101*, 1635.

(2) Goymer, P. *Nature Chem.* **2012**, *4*, 863.

(3) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740.

(4) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

(5) Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738.

(6) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 729.

(7) Desiraju, G. R. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 52.

(8) Arunan, E.; Desiraju, G. R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary, D. C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P. *Pure Appl. Chem.* **2011**, *83*, 1637.

(9) Zhang, P.; Chu, H.; Li, X.; Feng, W.; Deng, P.; Yuan, L.; Gong, B. *Org. Lett.* **2010**, *13*, 54.

(10) Cheng, C.-C.; Yen, Y.-C.; Chang, F.-C. *R. Soc. Chem. Adv.* **2011**, *1*, 1190.

(11) Mudraboyina, B. P.; Wisner, J. A. *Chem.--Eur. J.* **2012**, 14157.

(12) Li, J.; Pandelieva, A. T.; Rowley, C. N.; Woo, T. K.; Wisner, J. A. *Org. Lett.* **2012**.

(13) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *57*, 250.

(14) Kyogoku, Y.; Lord, R. C.; Rich, A. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1969**, *179*, 10.

(15) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.

(16) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.

(17) Sartorius, J.; Schneider, H.-J. *Chem.--Eur. J.* **1996**, *2*, 1446.

(18) Popelier, P. L. A.; Joubert, L. *J. Am. Chem. Soc.* **2002**, *124*, 8725.

(19) Quinn, J. R.; Zimmerman, S. C.; Del Bene, J. E.; Shavitt, I. *J. Am. Chem. Soc.* **2007**, *129*, 934.

(20) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.

(21) Murray, T. J.; Zimmerman, S. C.; Kolotuchin, S. V. *Tetrahedron* **1995**, *51*, 635.

(22) Murray, T. J.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1992**, *114*, 4010.

(23) Zimmerman, S.; Corbin, P.; Fuiita, M.; Springer Berlin / Heidelberg: 2000; Vol. 96, p 63.

(24) Kotera, M.; Lehn, J.-M.; Vigneron, J.-P. *J. Chem. Soc., Chem. Commun.* **1994**, 197.

(25) Gooch, A.; McGhee, A. M.; Renton, L. C.; Plante, J. P.; Lindsay, C. I.; Wilson, A. J. *Supramol. Chem.* **2009**, *21*, 12.

(26) Ośmiałowski, B.; Kolehmainen, E.; Kauppinen, R.; Kowalska, M. *Supramol. Chem.* **2011**, *23*, 579.

(27) Ośmiałowski, B.; Kolehmainen, E.; Gawinecki, R.; Dobosz, R.; Kauppinen, R. *J. Phys. Chem. A* **2010**, *114*, 12881.

(28) Ośmiałowski, B.; Kolehmainen, E.; Gawinecki, R.; Kauppinen, R.; Koivukorpi, J.; Valkonen, A. *Struct. Chem.* **2010**, *21*, 1061.

(29) Carboni, S.; Da Settimo, A.; Tonetti, I. *J. Heterocycl. Chem.* **1970**, *7*, 875.

(30) Caluwe, P.; Majewicz, T. G. *J. Org. Chem.* **1977**, *42*, 3410.

(31) Laxmi Madhavi, N. N.; Senthivel, P.; Nangia, A. *J. Phys. Org. Chem.* **1999**, *12*, 665.

(32) Bell, D. A.; Anslyn, E. V. *J. Org. Chem.* **1994**, *59*, 512.

(33) Bell, D. A.; Anslyn, E. V. *Tetrahedron* **1995**, *51*, 7161.

(34) Leigh, D. A.; Djurdjevic, S.; McNab, H.; Parsons, S.; Teobaldi, G.; Zerbetto, F. *J. Am. Chem. Soc.* **2007**, *129*, 476.

(35) Zimmerman, S. C.; Murray, T. J. *Tetrahedron Lett.* **1994**, *35*, 4077.

(36) Leigh, D. A.; Blight, B. A.; Camara-Campos, A.; Djurdjevic, S.; Kaller, M.; McMillan, F. M.; McNab, H.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2009**, *131*, 14116.

(37) Newman, S. G.; Taylor, A.; Boyd, R. J. *Chem. Phys. Lett.* **2008**, *450*, 210.

(38) Sijbesma, R. P.; Meijer, E. W. *Chem. Commun.* **2003**, 5.

(39) Schmuck, C.; Wienand, W. *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 4363.

(40) Beijer, F. H.; Kooijman, H.; Spek, A. L.; Sijbesma, R. P.; Meijer, E. W. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 75.

(41) Corbin, P. S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9710.

(42) Corbin, P. S.; Zimmerman, S. C.; Thiessen, P. A.; Hawryluk, N. A.; Murray, T. J. *J. Am. Chem. Soc.* **2001**, *123*, 10475.

(43) Ligthart, G. B. W. L.; Ohkawa, H.; Sijbesma, R. P.; Meijer, E. W. *J. Am. Chem. Soc.* **2004**, *127*, 810.

(44) Park, T.; Zimmerman, S. C.; Nakashima, S. *J. Am. Chem. Soc.* **2005**, *127*, 6520.

(45) Kuykendall, D. W.; Anderson, C. A.; Zimmerman, S. C. *Org. Lett.* **2009**, *11*, 61.

(46) Kuramoto, K.; Tarashima, N.; Hirama, Y.; Kikuchi, Y.; Minakawa, N.; Matsuda, A. *Chem. Commun.* **2011**, *47*, 10818.

(47) Li, Y.; Park, T.; Quansah, J. K.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2011**, *133*, 17118.

(48) Hisamatsu, Y.; Shirai, N.; Ikeda, S.; Odashima, K. *Org. Lett.* **2010**, *12*, 1776.

(49) Hua, Y.; Flood, A. H. *Chem. Soc. Rev.* **2010**, *39*, 1262.

(50) Brammer, S.; Lüning, U.; Kühl, C. *Eur. J. Org. Chem.* **2002**, *2002*, 4054.

(51) Taubitz, J.; Lüning, U. *Eur. J. Org. Chem.* **2008**, *2008*, 5922.

(52) Ligthart, G. B. W. L.; Guo, D.; Spek, A. L.; Kooijman, H.; Zuilhof, H.; Sijbesma, R. P. *J. Org. Chem.* **2007**, *73*, 111.

(53) Murray, T. J.; Zimmerman, S. C. *Tetrahedron Lett.* **1995**, *36*, 7627.

(54) Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. *Science* **1997**, *278*, 1601.

(55) Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. *J. Am. Chem. Soc.* **1998**, *120*, 6761.

(56) Hentschel, J.; Kushner, A. M.; Ziller, J.; Guan, Z. *Angew. Chem., Int. Ed. Engl.* **2012**, n/a.

(57) J. B. Folmer, B.; Cavini, E. *Chem. Commun.* **1998**, 1847.

(58) Folmer, B. J. B.; Sijbesma, R. P.; Versteegen, R. M.; van der Rijt, J. A. J.; Meijer, E. W. *Adv. Mater.* **2000**, *12*, 874.

(59) Eckelmann, J.; Dethlefs, C.; Brammer, S.; Doğan, A.; Uphoff, A.; Lüning, U. *Chem.--Eur. J.* **2012**, *18*, 8498.

(60) Han, J. T.; Lee, D. H.; Ryu, C. Y.; Cho, K. *J. Am. Chem. Soc.* **2004**, *126*, 4796.

(61) de Greef, T. F. A.; Meijer, E. W. *Nature* **2008**, *453*, 171.

(62) De Greef, T. F. A.; Smulders, M. M. J.; Wolffs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109*, 5687.

(63) Fox, J. D.; Rowan, S. J. *Macromolecules* **2009**, *42*, 6823.

(64) Brunsveld, L.; Folmer, B. J. B.; Meijer, E. W.; Sijbesma, R. P. *Chem. Rev.* **2001**, *101*, 4071.

(65) Aida, T.; Meijer, E. W.; Stupp, S. I. *Science* **2012**, *335*, 813.

(66) Wilson, A. J. *Soft Matter* **2007**, *3*, 409.

(67) Begall, S.; Cerveny, J.; Neef, J.; Vojtech, O.; Burda, H. *Proc. Natl. Acad. Sci. U. S. A.* **2008**.

(68) Sun, H.; Lee, H. H.; Blakey, I.; Dargaville, B.; Chirila, T. V.; Whittaker, A. K.; Smith, S. C. *The Journal of Physical Chemistry B* **2011**, *115*, 11053.

(69) Lafitte, V. r. G. H.; Aliev, A. E.; Horton, P. N.; Hursthouse, M. B.; Bala, K.; Golding, P.; Hailes, H. C. *J. Am. Chem. Soc.* **2006**, *128*, 6544.

(70) Greco, E.; Aliev, A. E.; Lafitte, V. G. H.; Bala, K.; Duncan, D.; Pilon, L.; Golding, P.; Hailes, H. C. *New J. Chem.* **2010**, *34*, 2634.

(71) Hisamatsu, Y.; Shirai, N.; Ikeda, S.-i.; Odashima, K. *Org. Lett.* **2009**, *11*, 4342.

(72) Li, X.; Fang, Y.; Deng, P.; Hu, J.; Li, T.; Feng, W.; Yuan, L. *Org. Lett.* **2011**, *13*, 4628.

(73) Taubitz, J.; Lüning, U. *Aust. J. Chem.* **2009**, *62*, 1550.

(74) Blight, B. A.; Hunter, C. A.; Leigh, D. A.; Mcnab, H.; Thomson, P. I. T. *Nature Chem.* **2011**, *3*, 244.

Chapter II

An AAAA-DDDD Quadruple Hydrogen Bond Array

Portions of this chapter have been published as: An AAAA-DDDD Quadruple Hydrogen Bond Array, Blight, B. A.; Hunter, C. A.; Leigh, D. A.; McNab, H.; Thomson, P. I. T. *Nature Chemistry* **2011**, *3*, 244-248.

Portions of this chapter will be published as: AAAA-DDDD quadruple hydrogen bond arrays featuring NH•••N and CH•••N hydrogen bonds, Leigh, D. A.; Robertson, C. C.; Slawin, A. M. Z.; Thomson, P. I. T. *Journal of the American Chemical Society,* submitted.

Acknowledgements

The following people are gratefully acknowledged for their contribution to this chapter: Professors Hamish McNab and David Leigh proposed the research, contributed to the writing of the paper and provided many useful discussions. Drs Barry Blight and Smilja Djurdjevic first synthesised and characterised compounds **5**, **6**, and **7** (References 13 & 14). Dr Barry Blight also contributed to the preparation of the original paper, and created Figures 2.5 and 2.7. Dr Scott Cockroft provided useful discussion and advice around the project. Prof. Christopher Hunter designed the complex stability models and calculations. (Section 2.5.2 and Equation 1). Prof. Alex Slawin carried out crystallographic analysis on **25** and **12•13**.

2.1 Synopsis

This chapter concerns the synthesis and subsequent analysis of the first definitively-realised contiguous quadruple hydrogen bonded array. The array utilises secondary electrostatic interactions between adjacent hydrogen bonds in order to maximise the stability, having all the hydrogen bond donors (D) on one component and all the hydrogen bond acceptors (A) on the other.

The AAAA-DDDD quadruple hydrogen bonding array exhibited exceptionally strong binding for a small-molecule hydrogen bonded complex. The complex was too strong to be accurately evaluated in the non-polar solvent CH_2Cl_2, and a series of competition experiments established a lower limit of $K_a > 3 \times 10^{12}$ M^{-1}, which corresponds to to a binding free energy (ΔG) in excess of -71 kJ mol^{-1} - more than 20% of the thermodynamic stability of a carbon-carbon covalent bond and higher than any other known small-molecule hydrogen bonded complex.

The binding was also evaluated in polar, competing solvents and found to be 1.5×10^6 M^{-1} in MeCN and 3.4×10^5 M^{-1} in 10% v/v DMSO/$CHCl_3$, almost 300 times stronger than the previous strongest small-molecule hydrogen bonded complex.

2.2 Introduction

Multipoint hydrogen bonding motifs are the cornerstone of recognition processes in biology and are increasingly featuring in the design of multifunctional materials and supramolecular polymers.[1-5] In 1990 Jorgensen suggested that secondary electrostatic interactions play an important role in the stability of arrays of contiguous hydrogen bonds (Figure 2.1).[6] A corollary of this is that having all the hydrogen bond donor groups (D) in one partner and all the hydrogen bond acceptor sites (A) in the other should be the arrangement that produces the strongest binding because all secondary electrostatic interactions between neighbouring hydrogen bond pairs are attractive. Although the significance of this effect has been questioned based on the results of quantum mechanical calculations on DNA base pairs,[7] it accounts well for the general experimental trends found for triple hydrogen bonded complexes (typical complex stability ADA-DAD (**1•2**) < AAD-DDA (**3•4**) < AAA-DDD (**5•6**) < AAA-DDD$^+$ (**7•6**))[8-14] and can be used quantitatively in empirical methods for predicting complex stabilities.[8,15]

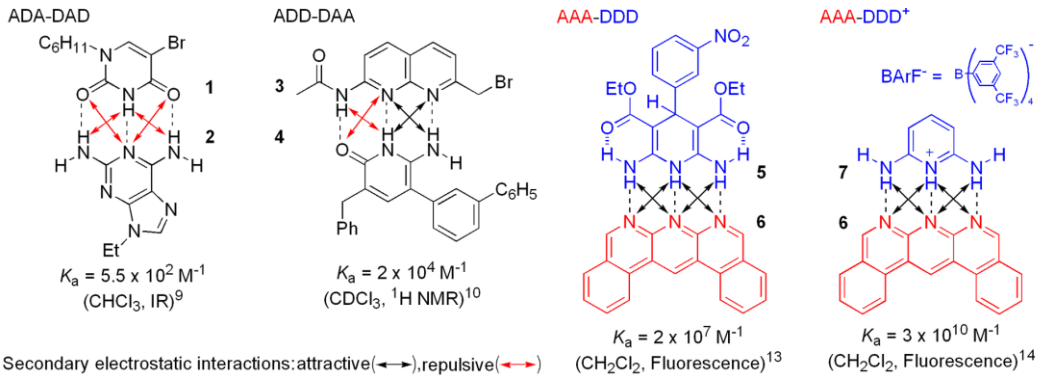


Figure 2.1 | Examples of triple hydrogen bond arrays and their association constants in various solvents. Different permutations of contiguous triple hydrogen bonded complexes ADA-DAD 1•2,[9] AAD-DDA 3•4,[10] AAA-DDD 5•6,[13] and cationic AAA-DDD$^+$ 7•6.[14] Arrows indicate secondary electrostatic interactions (black attractive; red repulsive).

The experimental trends are less clear-cut for quadruple hydrogen bond complexes (see, for example, Figure 2.2). However, there are far fewer examples of such arrays

and with a small sample set it is difficult to separate the contribution of the arrangement of the hydrogen bond pairs from other factors that contribute to complex stability, such as the hydrogen bond acidity/basicity of the functional groups, additional CH···O/N or multipole interactions, entropy effects, solvation, tautomerisation and the strength of homodimerisation of each partner. Indeed, of the six possible quadruple hydrogen bond permutations,[16-27] only four have been definitively experimentally realised to date: ADAD-DADA (e.g. **8•8**),[16] AADD-DDAA (e.g. **9•9**)[17-21] and ADDA-DAAD (e.g. **10•11**)[22-24] (Figure 2.2). The AAAD-DDDA motif remains experimentally unexplored and the AAAA-DDDD$^+$ motif **12•13** is presented in the current work.

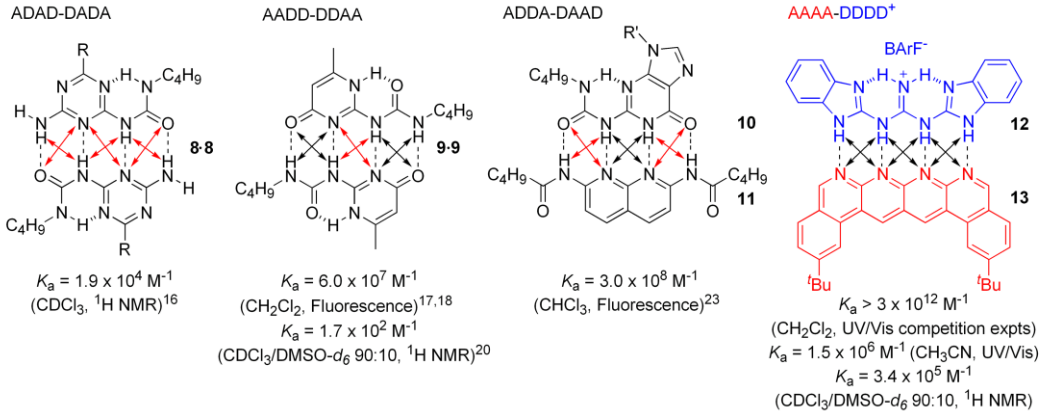
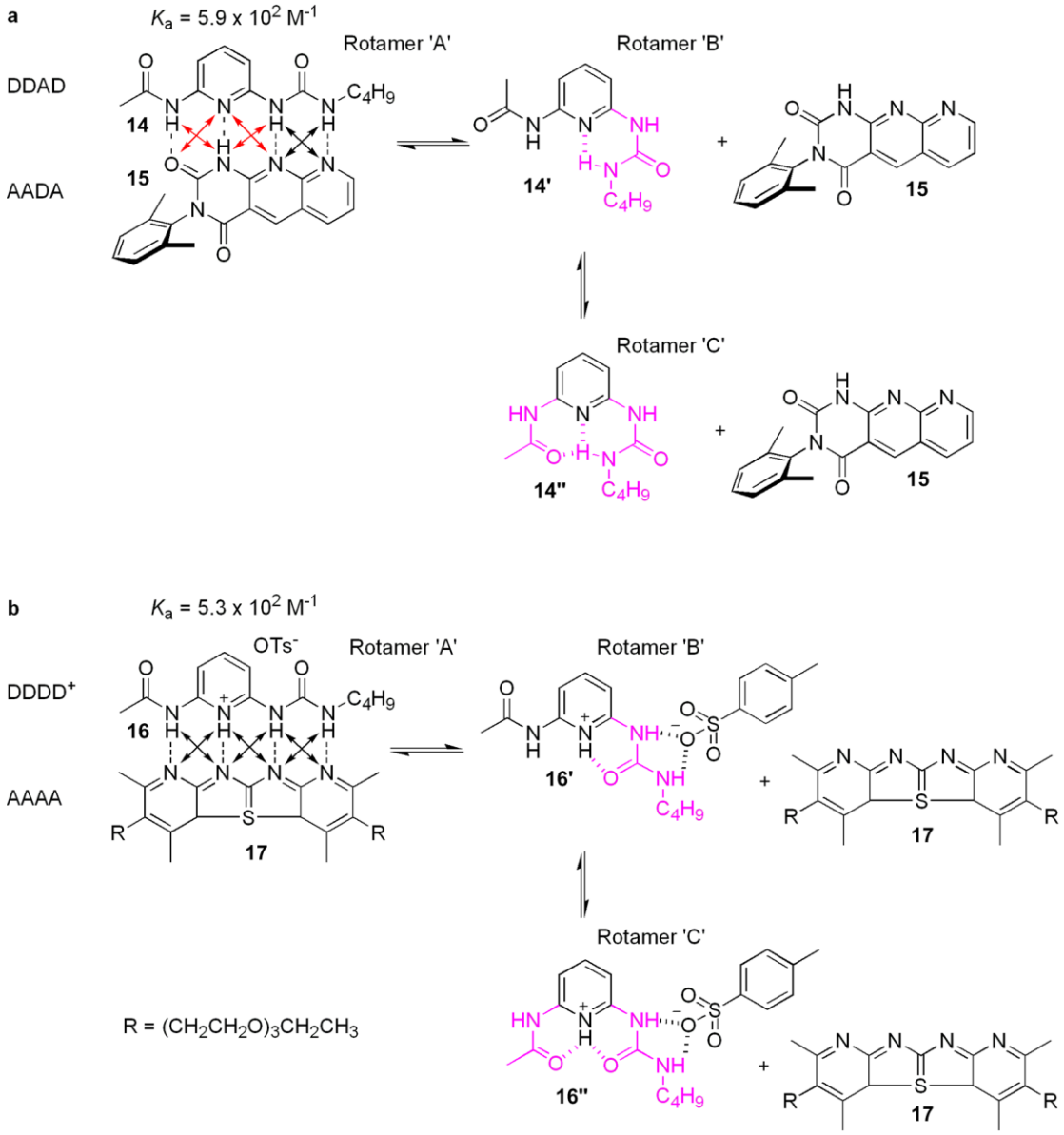


Figure 2.2 | Examples of quadruple hydrogen bond arrays and their association constants in various solvents. Different permutations of contiguous quadruple hydrogen bonded complexes: ADAD-DADA 8•8,[16] AADD-DAAD 9•9,[17,18,20] ADDA-DAAD 10•11[23] and AAAA-DDDD$^+$ 12•13 (this work). Arrows indicate secondary electrostatic interactions (black attractive; red repulsive).

Claims for an ADAA-DADD **14•15** and a cationic AAAA-DDDD$^+$ **16•17** complex have been made,[25,26] but both have anomalously low stability constants (K_a = 590 M^{-1} and K_a = 530 M^{-1} in $CDCl_3$, respectively), probably as a result of intramolecular hydrogen bonding that favours alternative conformations of **14** and **16** that do not correspond to the desired hydrogen bonding array. Additionally, competition from the strongly-coordinating tosylate anions could be further hindering the binding (Scheme 2.1, below)



Scheme 2.1 | Reported AADA-DDAD 14•15 (a, $K_a = 5.9 \times 10^2\ M^{-1}$) and AAAA-$DDDD^+$ 16•17 (b, $K_a = 5.3 \times 10^2\ M^{-1}$) complexes.[25,26] The DDAD and $DDDD^+$ components may favour rotamers that do not correspond to the intended H-bond motif.

Here we report on a readily-accessible AAAA-$DDDD^+$ complex **12•13** that exhibits exceptional complex stability even in hydrogen bond-disrupting solvent systems.

2.3 Design and synthesis of an AAAA-DDDD array

2.3.1 Design and synthesis of DDDD

A quadruple hydrogen bonded donor was initially sought using the same design principle as triple hydrogen bonded donor **7**. However, many potential designs would have suffered competitive folding (similar to those shown in scheme 2.1b) and so designs were sought which used preorganisation to eliminate favourable but non-binding foldamers. Figure 2.3 shows the evolution of proposed structures, towards the eventual synthesised design.

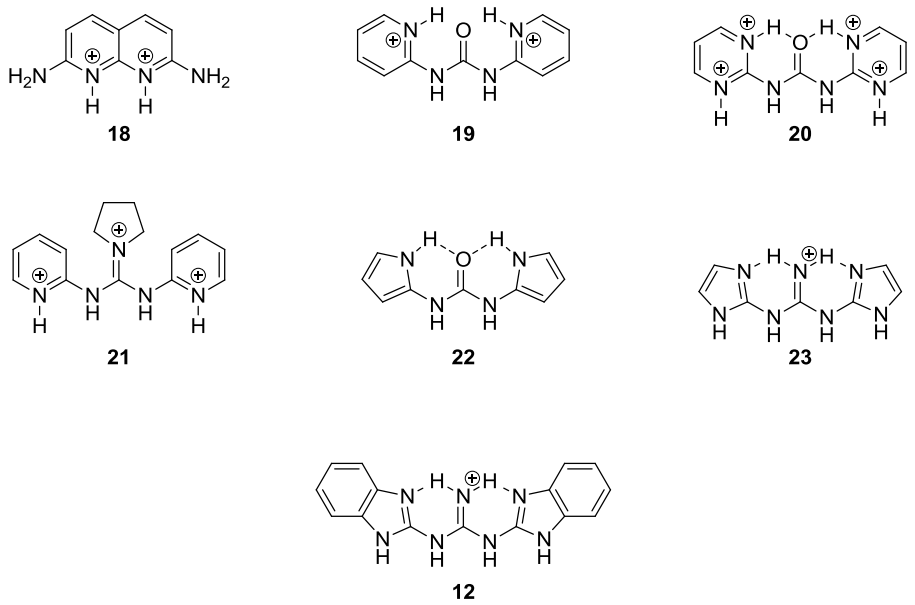
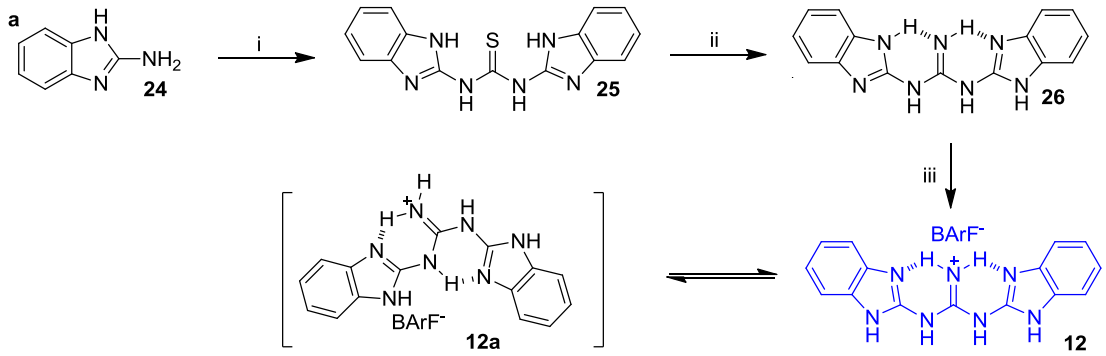


Figure 2.3 | Proposed designs of a quadruple hydrogen bond donor DDDD

2,6-Diaminopyridine can be protonated to convert it from a DAD to a DDD motif, switching an acceptor to a donor site (compound **7**, Figure 2.1). An attempt was made to protonate 2,7-diaminonaphthyridine twice to give **18**, but the precursor singly protonated species was not sufficiently basic or soluble for this to be possible. It was thought that a single protonation of two separate pyridine rings might be possible as in structure **19**, but it was concluded that this would only exist in the folded state as shown. Structure **20** would use two additional interactions as preorganisation, but

would bear four positive charges – a hindrance to solubility even with extremely lipophilic anions. Structure **21** would use an alkylated guanidinium to prevent internal hydrogen bonding, but again was thought to be too highly-charged to be feasible. Structure **22** was designed to use 5-membered heteroaromatic rings to present a neutral hydrogen bond donor, but would again suffer from internal hydrogen bonding. Structure **23** was then proposed, using internal hydrogen bonding to preorganise the presentation of a DDDD face. Availability of commercial starting material, ability to easily incorporate substituents, and ease of synthesis led to the closely-related bis-benzimidazole **12** being chosen as the target for synthetic work.

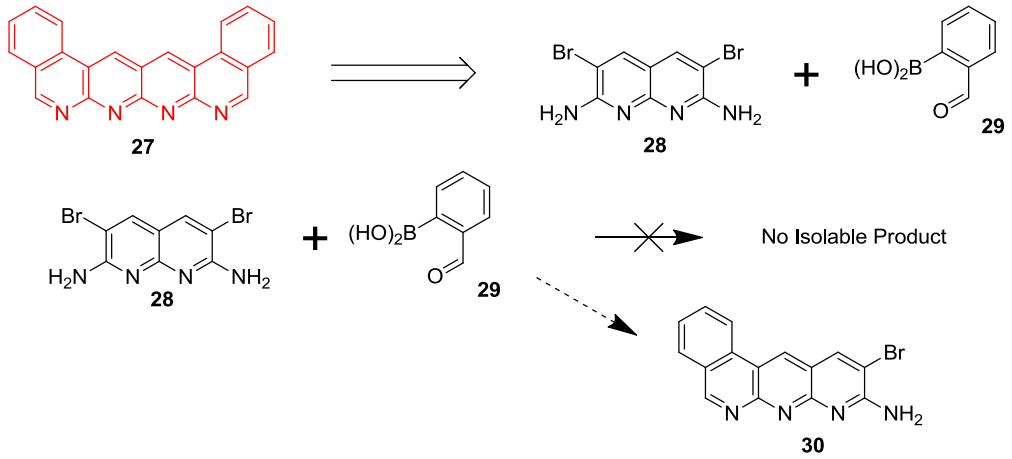
Compound **12** was synthesised from the commercially-available 2-aminobenzimidazole **24** as shown in Scheme 2.2. Reaction with carbon disulfide in pyridine at reflux gave thiourea **25**, which was converted to guanidine **26** *via* desulfuration to give a carbodiimide intermediate, which was quenched with ammonia. The guanidine **26** was readily protonated, then precipitated as the salt of the weakly-coordinating $[B(3,5\text{-}(CF_3)_2C_6H_3)_4]^-$ ($BArF^-$) anion.



Scheme 2.2 | Synthetic route to DDDD$^+$ 12. (i) CS_2, pyridine, 130 °C, 18 h, 81%; (ii) HgO, NH_3/MeOH, $CHCl_3$ 25 °C, 3 h, 50%; (iii) NaBArF, 8 M AcOH, 25 °C, 2 h, 40%.

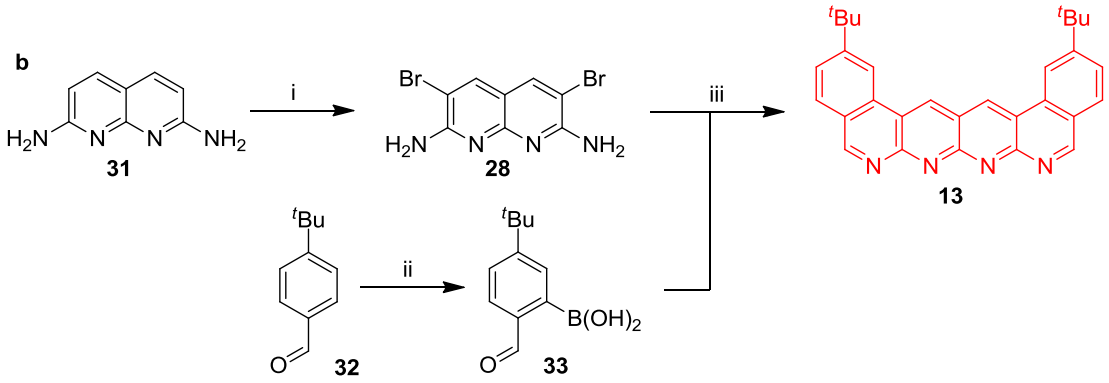
2.3.2 Design and synthesis of AAAA

A quadruple hydrogen bonded acceptor **27** was first designed using chemistry exactly analogous to that of the triple hydrogen bonded acceptor previously published (See Figure 1.11, Chapter I).[14] However, the first attempted cross-coupling between bromide **28** and boronic acid **29** was not successful under a variety of conditions (Scheme 2.3). Starting material was always consumed, but the reactions only produced material which was insoluble in non-polar solvents – potentially the mono-substituted derivative **30** which could have precipitated out, halting any further reaction.



Scheme 2.3 | Initial proposed route to an AAAA, and failed attempts at synthesis.

One of the desired characteristics of the target molecule was that it had to be soluble in CH_2Cl_2 and $CHCl_3$ in order to carry out binding constant determinations, so ^{t}Bu groups were chosen to modify the lead AAAA structure **27** in order to disrupt intramolecular π-stacking and thus enhance solubility.



Scheme 2.4 | Synthetic route to quadruple hydrogen bond acceptor 13. (i) NBS, DMF, -10 to 25 ºC, 3 h, 60%; (ii) $NHCH_3(CH_2CH_2)N(CH_3)_2$, BuLi then $B(OMe)_3$, THF, -78 to 25 ºC, 16 h, 36%; (iii) Na_2CO_3, H_2O/DME (1:1), $Pd(PPh_3)_4$, 80 ºC, 1.5 h, 27%.

The corresponding synthesis of **13** was successful and **12** was assembled in acceptable yield through bromination of known 2,7-diamino-1,8-naphthyridine **31** to the bis-bromide **28**, followed by a one-pot double Suzuki coupling-cyclisation-aromatisation procedure[13,14] with boronic acid **33** (Scheme 2.4). Compound **33** itself could be synthesised from the commercially available aldehyde **32** *via* a one-pot procedure in which trimethyl ethylene diamine was used simultaneously as an *in-situ* protecting and directing group, before being cleaved in the workup to return the desired aldehyde functionality.

2.4 Evidence of formation of an AAAA-DDDD array

2.4.1 Evidence of formation of an AAAA•DDDD complex by NMR

Before evaluating the binding constant of **13•12**, studies were carried out to establish that a specific, directional, intramolecular hydrogen bonding interaction was happening. In order to ensure that experiments evaluating the binding strength of **13** were not complicated by additional equilibrium processes, we also evaluated the strength of **13•13** dimerisation. An initial solution of a known quantity of **13** in $CDCl_3$ was successively diluted, and the chemical shift of the resonance at 9.78 ppm monitored*. The results are shown in Figure 2.4. The linear shape of the graph indicates that the dimerisation constant $K_{dim} < 100\ M^{-1}$. To confirm this, the general equation for ^1H NMR dimerisation[29] was modelled with parameters from our experiment at a variety of binding constants (K_{est}) and the results are also shown in Figure 2.4. Significant deviation from linearity is evident for $K_{est} > 100\ M^{-1}$. In order to accurately quantify K_{dim}, higher concentrations would need to be utilised, however the limited solubility of **13** in $CDCl_3$ prevented this.

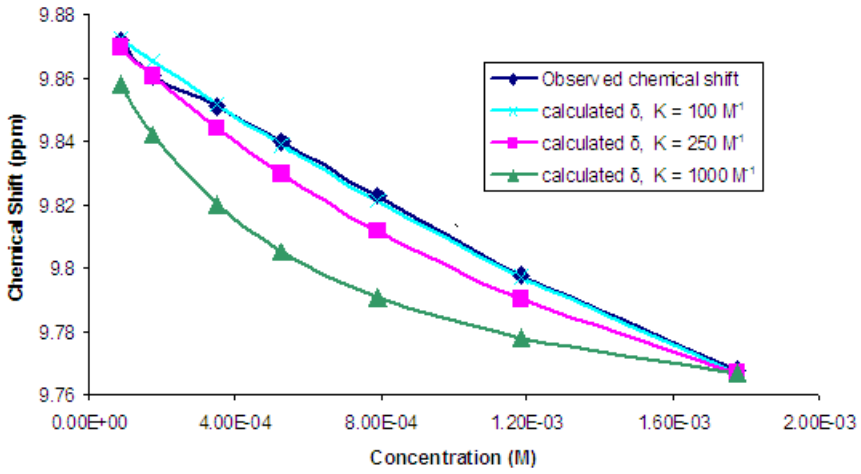


Figure 2.4 | Observed (for 13•13; blue) and theoretical (turquoise, pink, green) chemical shifts for various dimerisation constants.

* This resonance displayed the largest change in chemical shift, but other protons were similar.

The upfield shift of the signal at higher concentrations suggests that π-π-stacking may be occurring, however, at the concentrations used in the heterocomplexation experiments the presence of **13•13** is negligible.

The intramolecular **12•13** complex was then studied by ^{1}H NMR spectroscopy (Figure 2.5). Particularly noteworthy are the large downfield shifts (up to 10 ppm) of the NH protons of **12** upon complex formation with **13**, and the upfield shifts of the benzimidazole CH protons as the NH bonds become more polarised through hydrogen bonding.

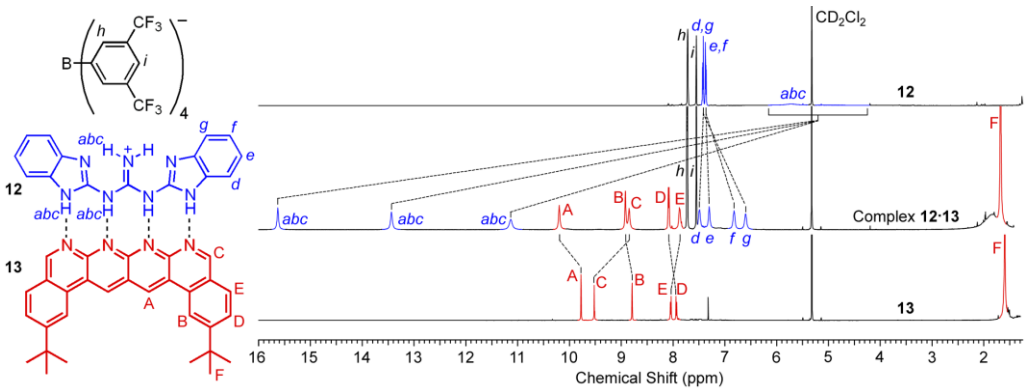


Figure 2.5 | ^{1}H NMR Spectra (1 mM, 500 MHz, CD_2Cl_2, 298 K) of 12 (top), complex 12•13 (middle), and 13 (bottom). Dashed lines show the changes in chemical shift of the resonances in 12 and 13 upon formation of complex 12•13. Used with permission of Dr. Barry A. Blight.

Protons *abc* were in chemical exchange and could not be distinguished. Protons *d/g* and *e/f* are equivalent in the spectrum of free **12**, probably due to fast conversion between two rotamers (**12** and **12a**, Scheme 2.2).[28] ROESY experiments showed through-space interactions between the H_C protons of **13** and the H_d protons of **12** (Figure 2.6).

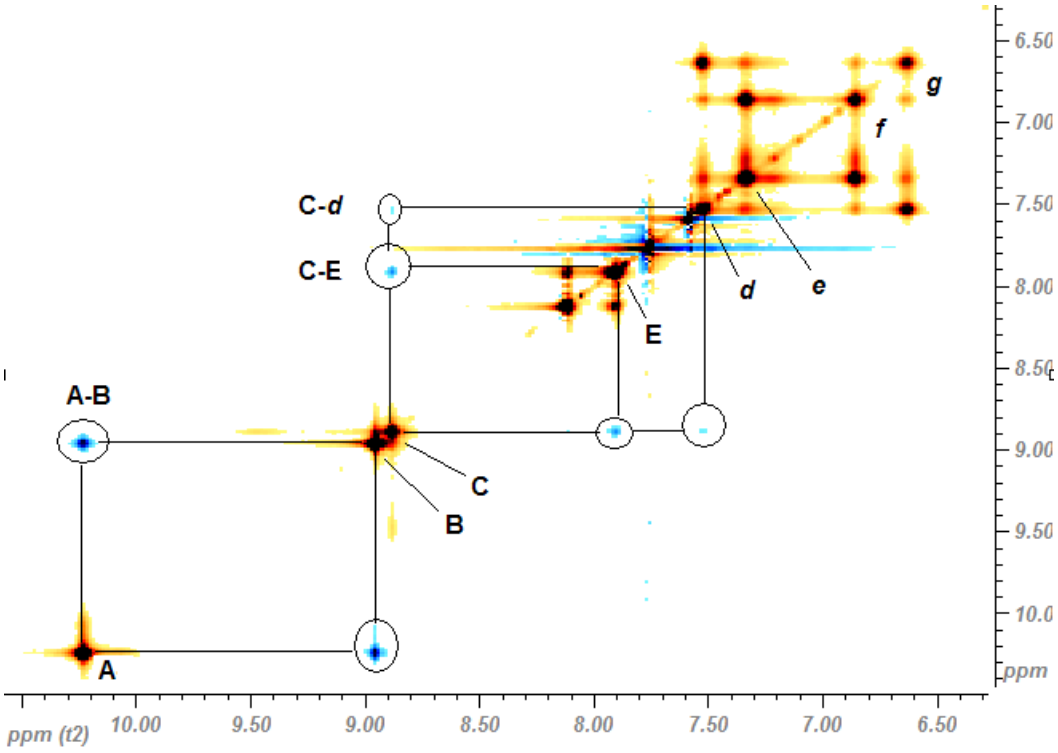
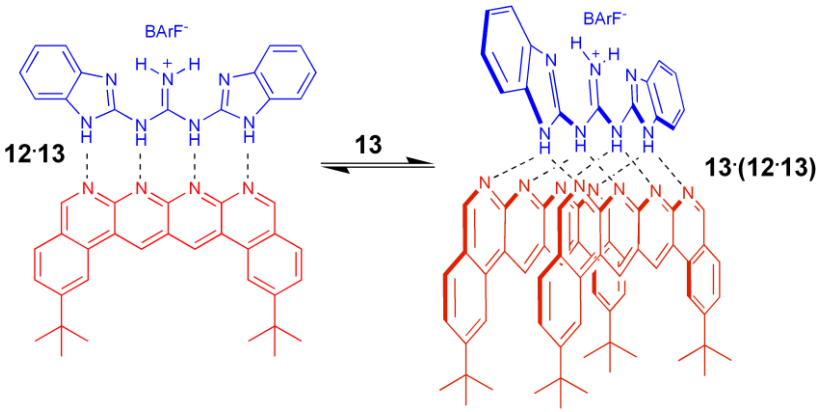


Figure 2.6 | Partial ^{1}H-^{1}H ROESY NMR spectrum (1mM, 500 MHz, CD_2Cl_2, 298K) of 12•13 illustrating a through space interaction between 12 and 13 (C-d), also enabling the assignment of d, e, f, and g.

An additional equilibrium between **12** and two equivalents of **13** was observed (see Section 2.5), and it was proposed that this was due to an arrangement of bifurcated hydrogen bonds (Scheme 2.5). The steric repulsion of **13** with itself may be offset by a staggered arrangement of ^{t}Bu groups, and the weak self-association of **13** (Figure 2.4) will also make this process more favourable. Similar ^{1}H NMR shifts were observed (Figure 2.7) as to the 1:1 equilibrium (Figure 2.5).



Scheme 2.5 | Possible structure of 13·(12·13) in equilibrium with 12·13 in the presence of excess 13.

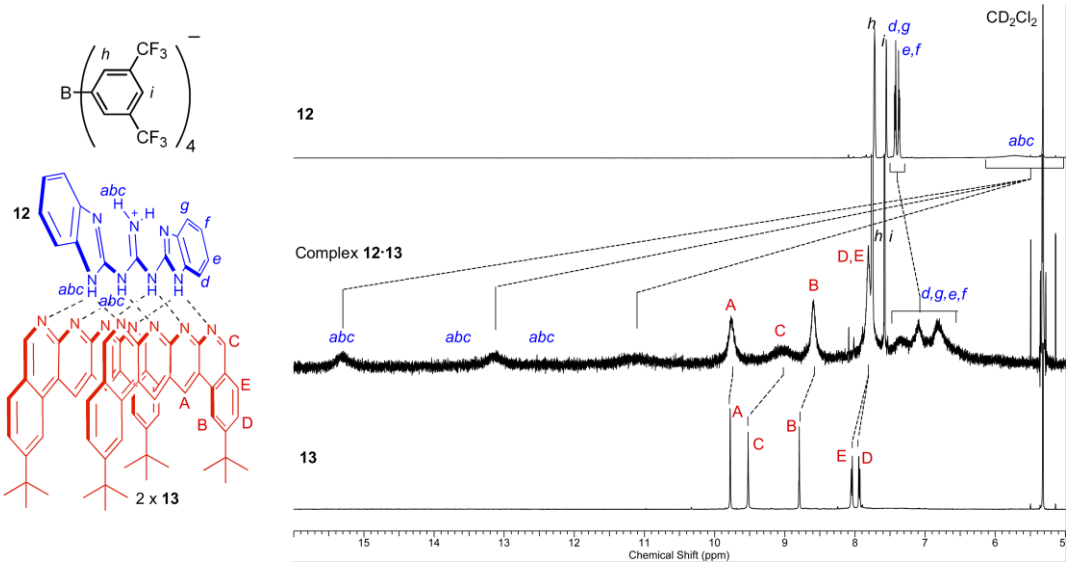
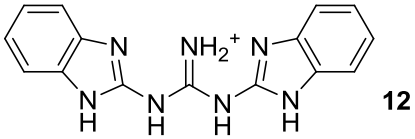


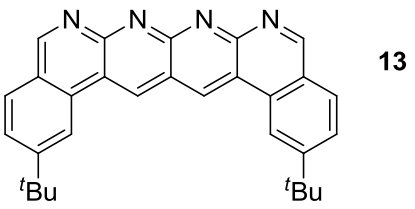
Figure 2.7 | ^{1}H NMR spectra (500 MHz, CD_2Cl_2, 298K) of 12 (top), 13·(12·13) (middle), and 13 (bottom) showing the changes to 12 and 13 upon complexation to form 13·(12·13). This experiment simulates an NMR titration at the point of adding 0.5 equivalents of 12 to 13 before reaching a 1:1 stoichiometry. Comparing to Figure 2.5, Protons B and D of 13 shift upfield as the above 13·(12·13) complexed is reached before shifting back downfield as the stoichiometry of the components becomes 1:1 (e.g. 12·13 in Fig. 2.5). Used with permission of Dr. Barry A. Blight.

2.4.2 Evidence of formation of an AAAA•DDDD complex by Mass Spectrometry

An ESI-MS experiment performed on a 1:1 mixture of **12·13** gave a peak corresponding to the 1:1 complex, which is positively charged and has a mass of 736.4 (Figure 2.9). Also present are signals for $\mathbf{12}^{+}$ (292), $\mathbf{12}\cdot 5MeCN\cdot Na^{+}$ (519), and $\mathbf{13}\cdot\mathbf{H}^{+}$ (445). The isotopic pattern (Figure 2.10) is consistent with theory (Figure 2.8), having an M+1 of *ca*. 50% intensity. Only the counter-ion was observed in negative ion detection mode. An accurate mass determination was not attempted by the time of preparation of the current work.

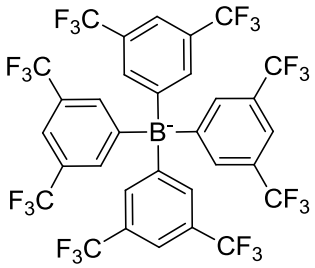


12



13

Chemical Formula: $C_{45}H_{42}N_{11}^{+}$
Exact Mass: 736.36
m/z: 736.36 (100.0%), 737.37 (49.2%), 738.37 (11.8%), 737.36 (4.1%), 739.37 (2.3%), 738.36 (2.1%)



Chemical Formula: $C_{32}H_{12}BF_{24}^{-}$
Exact Mass: 863.07
m/z: 863.06 (100.0%), 864.07 (34.7%), 862.07 (24.8%), 863.07 (8.6%), 865.07 (5.8%), 864.08 (1.5%)

Figure 2.8 | Predicted isotopic distribution patterns for 12•13 and the BArF- counterion.

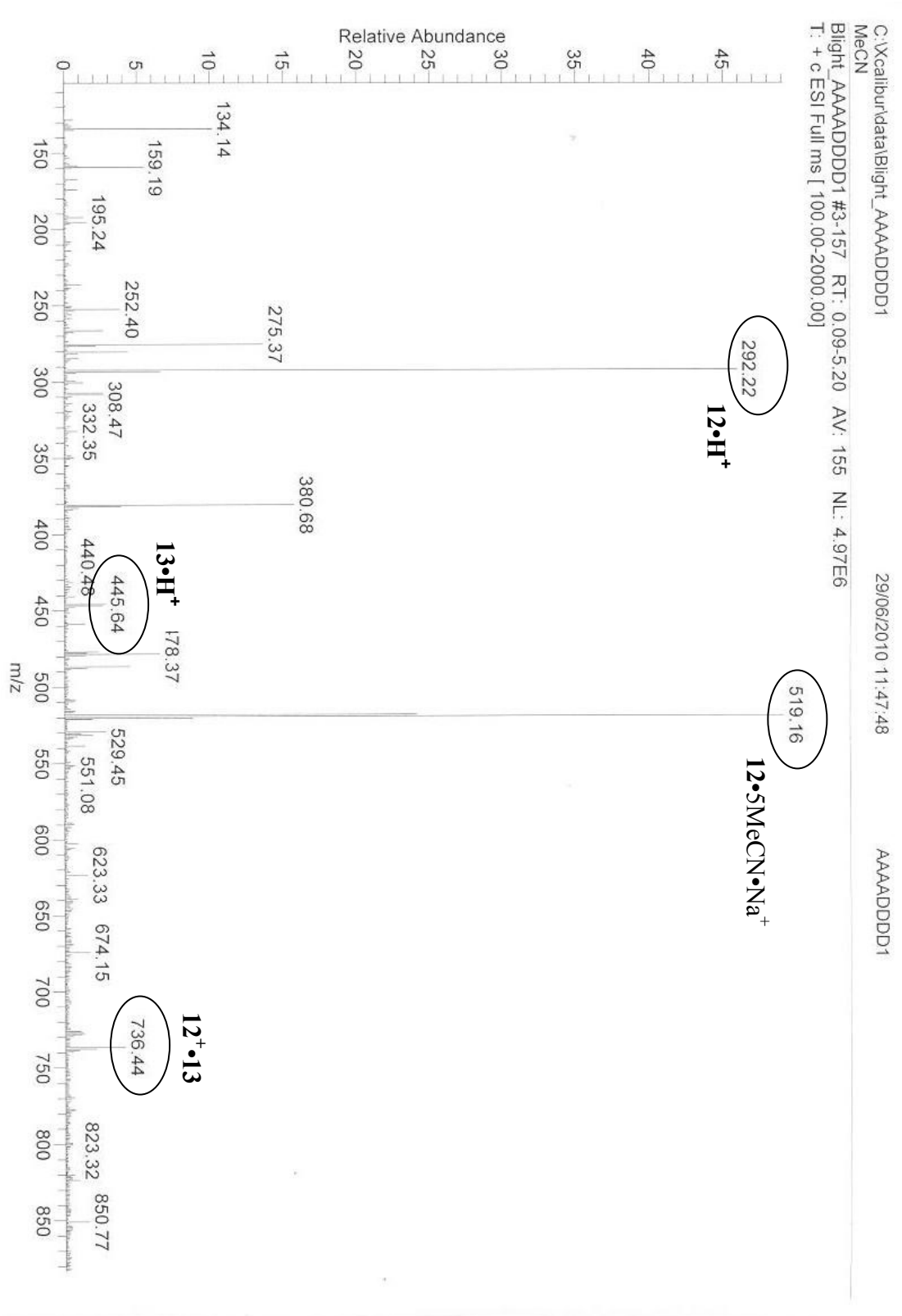


Figure 2.9 | ESI+ mass spectrum of 12•13 in an MeCN/CH_2Cl_2 matrix, m/z of 100-900.

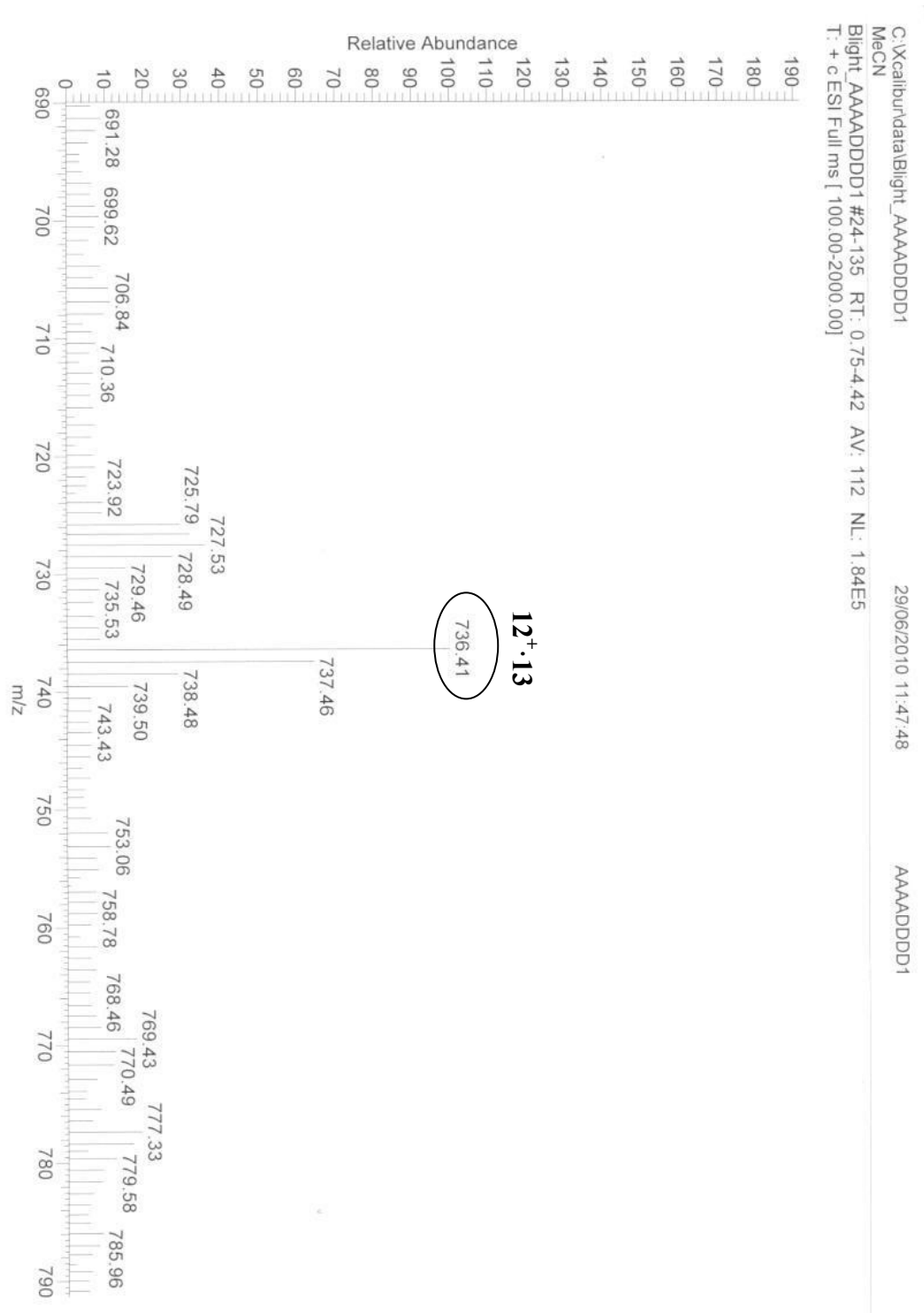


Figure 2.10 | ESI+ mass spectrum of 12·13 in an MeCN/CH_2Cl_2 matrix, m/z in the range of 690-790 showing the peak corresponding to the complex, with correct isotopic distribution pattern.

2.4.3 Evidence of formation of an AAAA•DDDD complex by X-ray Diffraction

A single crystal of **12•13** suitable for X-ray diffraction was grown by slow diffusion of hexane vapour into a solution of **12•13** in CH_2Cl_2. The X-ray crystal structure (Figure 2.11) shows that the conformation of **12** is locked by two intramolecular hydrogen bonds that present the four N-H hydrogen bond donors along one edge of the molecule. The hydrogen bonding edge of **6** shows a slight curve in the solid state, with the outer pyridine rings closer to the donor array than the inner pyridine rings. The average peripheral NH•••N distance is 1.79 Å, 0.17 Å shorter than the inner NH•••N distance (1.96 Å) (Figure 2.11a). The four NH•••N hydrogen bonds are all close to linear, in contrast to the staggered arrangement observed in the X-ray structure of an AAA-DDD hydrogen bond array **7•6** previously reported by our group.[14] The phenyl groups of **13** are slightly twisted away from the plane formed by the fused pyridine rings (Figure 2.11b).

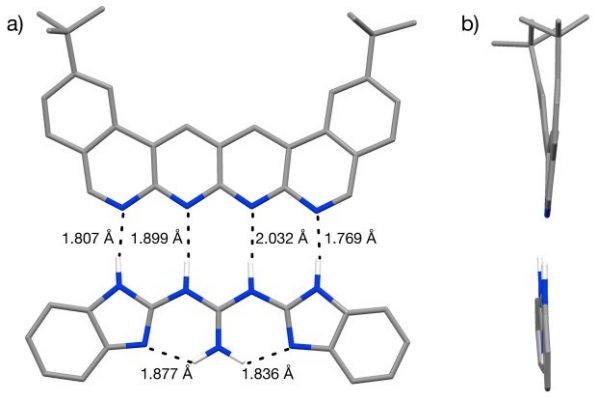


Figure 2.11 | X-Ray crystal structure of 12•13 from a single crystal grown from slow diffusion of hexane into a saturated solution of the complex in CH_2Cl_2 (see Section 2.7.6 for crystallographic details). Carbon atoms are shown in grey, nitrogen atoms are blue and selected hydrogen atoms are white. Counter-anion omitted for clarity. (a) View 'face on' to the aromatic rings showing the linearity of the hydrogen bonds in the AAAA-DDDD array. (b) View along the edge of the complex showing the planarity of 12 and the slight twist of the pendant phenyl rings in 13.

2.5 Evaluation of binding of an AAAA-DDDD array

2.5.1 Evaluation of an AAAA•DDDD complex in CH_2Cl_2

UV/vis titration of **13** with **12** in CH_2Cl_2 (~5×10^{-5} M) showed a decrease in the absorption spectrum of **13** (λ_{max} = 426 nm) accompanied by the appearance of a new species bathochromically shifted by 11 nm (Figure 2.12). When data fitting was attempted to determine K_a, only a lower limit of 10^5 M^{-1} could be established.

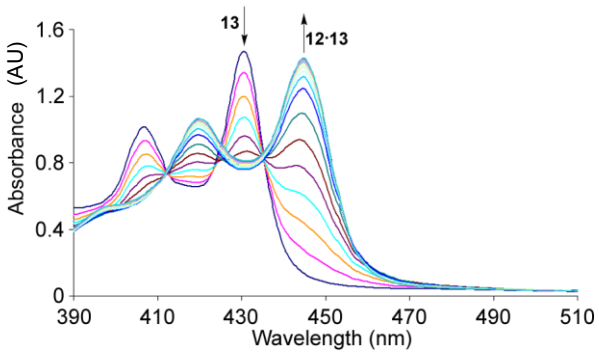
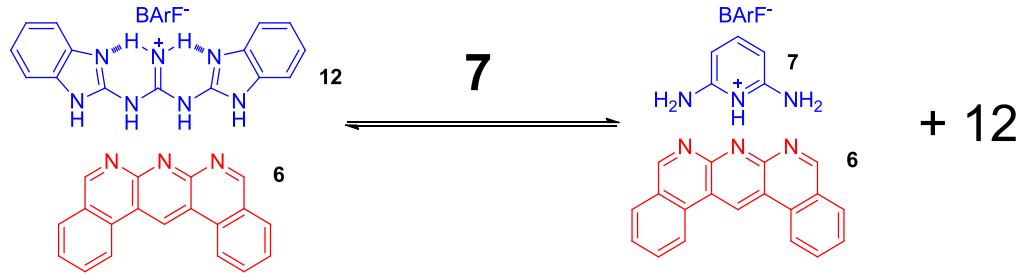


Figure 2.12 | UV/vis titration of AAAA 13 with DDDD$^+$ 12 in CH_2Cl_2. UV/vis spectra of 13 (ca. 5×10^{-5} M) upon addition of 12 (0 → 5 equiv), maintaining the concentration of 13 constant, in CH_2Cl_2 at 298 K.

In order to measure a high binding constant directly, experiments run at higher dilution would normally be carried out. However, compound **13** proved photochemically unstable at the light intensities required to measure changes in fluorescence at low concentrations (ca. 10^{-10} M). Therefore competition experiments were performed in CH_2Cl_2 to compare the association constant for **12•13** to that of known complex AAA-DDD$^+$ **7•6**. The binding constant in CH_2Cl_2 of **7·6** was known from previous fluorescence titration measurements,[14] and competitive binding between **7** and **12** would therefore allow the determination of $K_{12\cdot6}$ (Scheme 2.6).



Scheme 2.6 | Addition of 7 will displace 12 from 6. Changes in the UV/vis spectrum of 6 can be monitored without any overlap from 12 or 7.

Complex **7•6** was titrated with **12** (followed by changes in the UV/vis spectrum) to give a $K_a = 1 \times 10^{10}$ M^{-1} (ΔG = -57.1 kJ mol^{-1}) for the mismatched AAA-DDDD$^+$ complex **12•6**, slightly less than that of **7•6** ($K_a = 3 \times 10^{10}$ M^{-1})

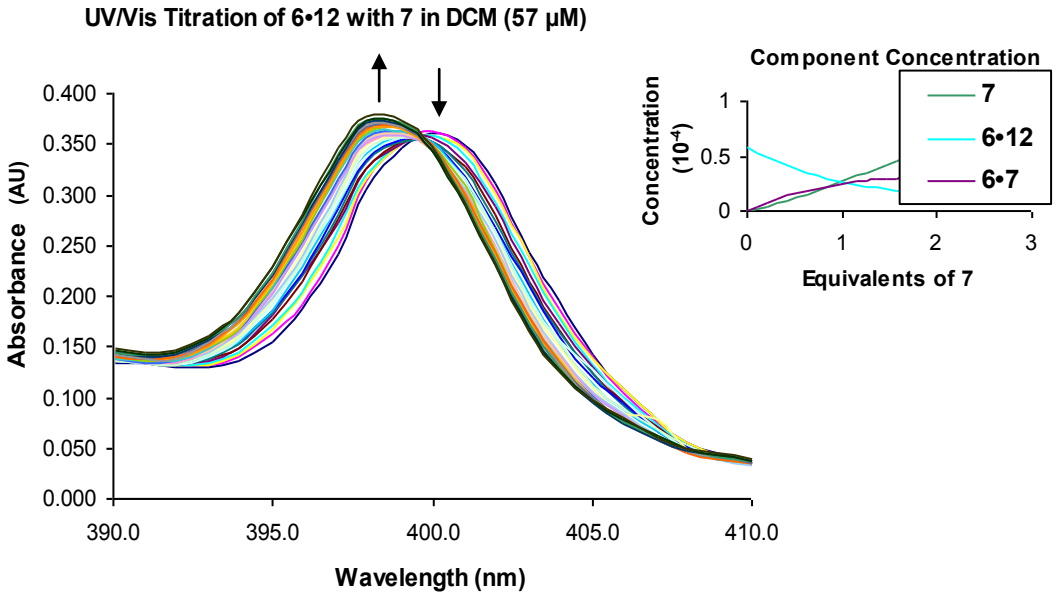


Figure 2.13 | UV/vis spectra of 6•12 (5.7×10^{-5} M) upon addition of 7 (0 → 3 equiv), while maintaining the concentration of 6•12 constant, in CH_2Cl_2 at 298 K. Component distribution over the course of the titration experiment is also illustrated (inset). Details of experimental and data fitting procedures in Section 2.7

The resulting spectra (Figure 2.13) showed a hypochromic shift of approximately 2 nm, corresponding to the difference between **12·6** and **7·6**. The shift displayed an isosbestic point, implying a simple 1:1 equilibrium between **12·6** and **7·6** with no

ternary complex.[*] The data was subjected to multivariate analysis, which used data at every wavelength simultaneously to produce a best fit value for the binding constant and deconvoluted spectra of **12•6** and **7•6**, all at the same time (Figure 2.14). Repeating experiments in triplicate yielded a binding constant of $K_a = 1 \times 10^{10} \pm 1.6 \times 10^9$ M^{-1}.

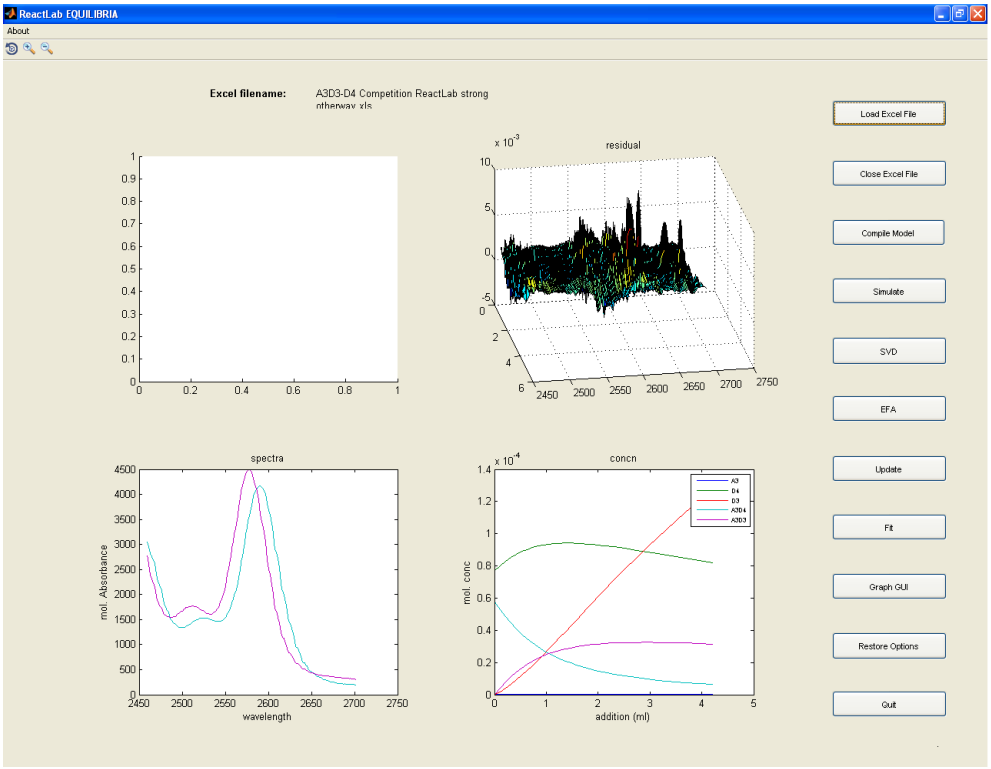


Figure 2.14 | UV/vis titration of 6•12 with 7 in CH_2Cl_2, ReactLab Working Window. Top left shows a plot of the convergence parameter while a fit is being calculated. Top right, a 3D plot of the residuals for the fit. Bottom left, the calculated spectrum for each isolated species. Bottom right, concentration profile for the species in solution.

A UV/vis titration of **12•6** with **13** gave a lower limit for the K_a of **12•13** of 10^{11} M^{-1} (the experiment being complicated by 1:1 and 2:1 complexes of **12•13** and the overlapping absorption spectra of **6** and **13**). The reverse experiment was performed (titrating AAAA-DDDD$^+$ **12•13** with a large excess of AAA **6** to liberate **13**) in order

[*] The absence of an isosbestic point rules out a simple 1:1 equlibrium but not the converse.[41] However, higher-order equilibrium containing multiple equivalents of **12** or **7** can be ruled out due to charge repulsion, and equilibrium containing multiple equivalents of **6** could be successfully identified when they occurred in other cases.

to determine the stability of **12•13** at high concentrations of **6** (Figure 2.15). Free **13** began appearing above a threshold of 350 equivalents of **6**, indicating a $K_a > 3 \times 10^{12}$ M^{-1} ($\Delta G < -71$ kJ mol^{-1}) for AAAA-DDDD$^+$ **12•13** in CH_2Cl_2.

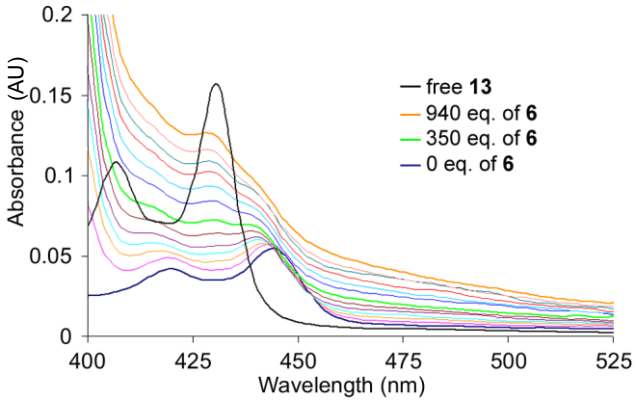


Figure 2.15 | UV/vis competition experiment in which 13 is displaced from 12•13 by a large excess of 6. UV/vis spectra of 12•13 (ca. 2 × 10^{-7} M) upon addition of 6 (0 → 950 equiv, 50 equiv aliquots), in CH_2Cl_2 at 298 K. Spectra at 0 equivalents (blue), 350 equivalents (green), and 940 equivalents (orange) of added 6 are shown with thicker linewidths for clarity. For comparison, the spectrum of free 13 (black) is also shown. Details of experimental and data fitting procedures in Section 2.7

2.5.2 Evaluation of an AAAA•DDDD complex in MeCN

Solvent competition for hydrogen bonding sites can dramatically influence complex stability. If the thermodynamic contributions of individual hydrogen bonds in a multiply hydrogen bonded complex are assumed to be additive in free energy, the magnitude of the expected solvent effects can be estimated using Hunter's hydrogen bond parameters[30] in conjunction with Equation 1:

$$\Delta G^0 = -\sum_i (\alpha_i - \alpha_S)(\beta_i - \beta_S) + 6 \text{ kJ mol}^{-1} \qquad \text{Eq. 1}$$

where α_S and β_S are the hydrogen bond parameters of the solvent, α_i and β_i are the hydrogen bond parameters for the solute interaction sites, the sum is over all solute-solute interactions and 6 kJ mol^{-1} is a constant representing the cost of forming a bimolecular complex in solution.

If we assume that all the hydrogen bond donor sites on **12** and **7** are similar and comparable to other ammonium and guanidinium cation hydrogen bond donors that have been characterised experimentally,[31] it is possible to estimate stability constants for the **7•6** and **12•13** complexes in CH_2Cl_2 and in MeCN from Equation 1. The calculated values in CH_2Cl_2, 4×10^9 M^{-1} and 2×10^{13} M^{-1} for three and four hydrogen bonds respectively, are consistent with the experimental measurements. Changing the solvent to MeCN, a solvent in which the strength of solute-solute interactions are significantly reduced,[30] lowers the stability constants estimated using Equation 1 to 2×10^3 M^{-1} and 4×10^4 M^{-1}, respectively.

12 and **13** did not give good NMR spectra in CD_3CN, so NMR titration could not be used to determine the binding constant experimentally. Isothermal Titration Microcalorimetry (ITC) was initially attempted in collaboration with Alan Cooper (Glasgow). ITC is a technique that measures binding constants by the heat change when a binding event happens, not relying on any particular spectroscopic method. However, individual titrations require upwards of several hours to conduct, and the choice of solvent is limited. Additionally, data modelling requires deconvolution of

enthalpic and entropic contributions, and is complicated by the presence of multiple equilibria. We were never able to obtain repeatable data and the time cost was prohibitive.

UV/vis titration of AAA-DDD$^+$ complex **7•6** gave a $K_a = 5.4 \times 10^3$ M^{-1} (ΔG = -21.3 kJ mol^{-1}) in MeCN, far less than the $K_a = 3.0 \times 10^{10}$ M^{-1} in CH_2Cl_2 and in good agreement with the value estimated from Equation 1. During the titration of **13** with **12** (Figure 2.16), in addition to the 1:1 complex a 2:1 binding mode was also observed, and the binding energy of both modes could be determined simultaneously by choice of data fitting parameters. The complex stabilities of **12·13** and **13•(12•13)** were determined to be $K_{\mathbf{12•13}} = 1.5 \times 10^6$ M^{-1} (ΔG = -35.2 kJ mol^{-1}) – 280x stronger than **7•6** and 50x stronger than the value estimated from Equation 1 - and $K_{\mathbf{13•(12•13)}} = 2.4 \times 10^5$ M^{-1} (ΔG = -30.6 kJ mol^{-1}), respectively.

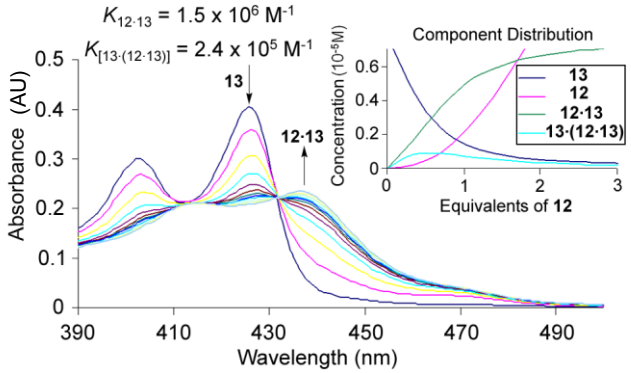


Figure 2.16 | UV/vis titration experiment in which $K_{12•13}$ and $K_{[13•(12•13)]}$ are measured by addition of 12 to 13 in MeCN. UV/vis spectra of 13 (ca. 8 × 10^{-6} M) upon addition of 12 (0 → 5 equiv), while maintaining the concentration of 13 constant, in MeCN at 298 K. Component distribution over the course of the titration experiment is also illustrated (inset). Association constants for 12•13, which were modelled with a 2:1 equilibrium, are $K_{12•13} = 1.5 \times 10^6$ M^{-1} (ΔG = -35.2 kJ mol^{-1}) for 12•13 and $K_{13·(12·13)} = 2.4 \times 10^5$ M^{-1} (ΔG = -30.6 kJ mol^{-1}) for the binding of a second AAAA unit to this complex to form 13•(12•13). Details of experimental and data fitting procedures in Section 2.7

The formation of **13•(12•13)** is probably due to favorable electrostatic interactions between **13** and **12•13**, and favourable pi-stacking between different molecules of **13** (Figure 2.4). No other higher-order equilibria were observed in any other complexes studied in MeCN, as confirmed by Job Plots (see Section 2.7.4 for experimental details).)

2.5.3 Evaluation of an AAAA•DDDD complex in $CHCl_3$/DMSO mixtures

MeCN is a stronger hydrogen bond acceptor than CH_2Cl_2, but a weaker acceptor than a pyridine nitrogen (the basic acceptor unit present in **13**), and so the moderation of the stability of the AAAA-DDDD$^+$ complex is as expected. Dimethylsulfoxide (DMSO) is an even stronger hydrogen bond acceptor than pyridine and so might be anticipated to eliminate binding entirely. This is the case in pure DMSO, but the AAAA-DDDD$^+$ complex proved remarkably stable in a solution of 10% DMSO in $CHCl_3$ ($K_{\mathbf{12•13}}$= 3.4 × 10^5 M^{-1} (ΔG = -31.6 kJ mol^{-1}), $K_{\mathbf{13•(12•13)}}$ = 1.4 × 10^5 M^{-1} (ΔG = -29.4 kJ mol^{-1}), despite being held together by only four intercomponent NH···N hydrogen bonds (see Section 2.7.5 for experimental details). The UV/vis spectra of these titrations (Figure 2.17) exhibited non-isosbestic points indicating a non-1:1 equilibrium, which was verified by ^{1}H NMR (Figure 2.18) and a Job Plot (Figure 2.19).

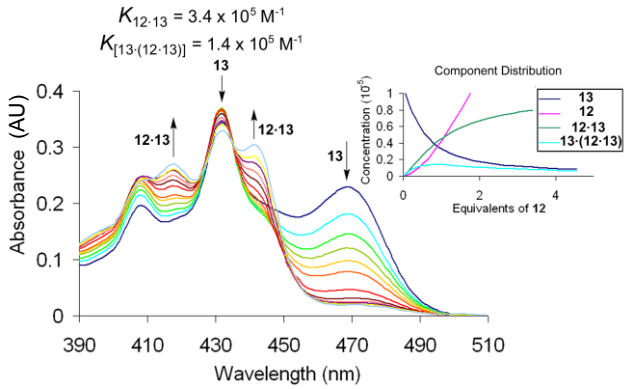


Figure 2.17 | UV/vis titration experiment in which $K_{12•13}$ and $K_{[13•(12•13)]}$ are measured by addition of 12 to 13 in a solution of 10% DMSO in $CHCl_3$. UV/vis spectra of 13 (ca. 1 × 10^{-5} M) upon addition of 12 (0 → 4.5 equiv), while maintaining the concentration of 13 constant, in a solution of 10% DMSO in $CHCl_3$ at 298 K. Component distribution over the course of the titration experiment is also illustrated (inset). Association constants for 12•13, which were modelled with a 2:1 equilibrium, are $K_{12•13}$ = 3.4 × 10^5 M^{-1} (ΔG = -31.6 kJ mol^{-1}) for 12•13 and $K_{13·(12·13)}$ = 1.4 × 10^5 M^{-1} (ΔG = -29.4 kJ mol^{-1}) for the binding of a second AAAA unit to this complex to form 13•(12•13).

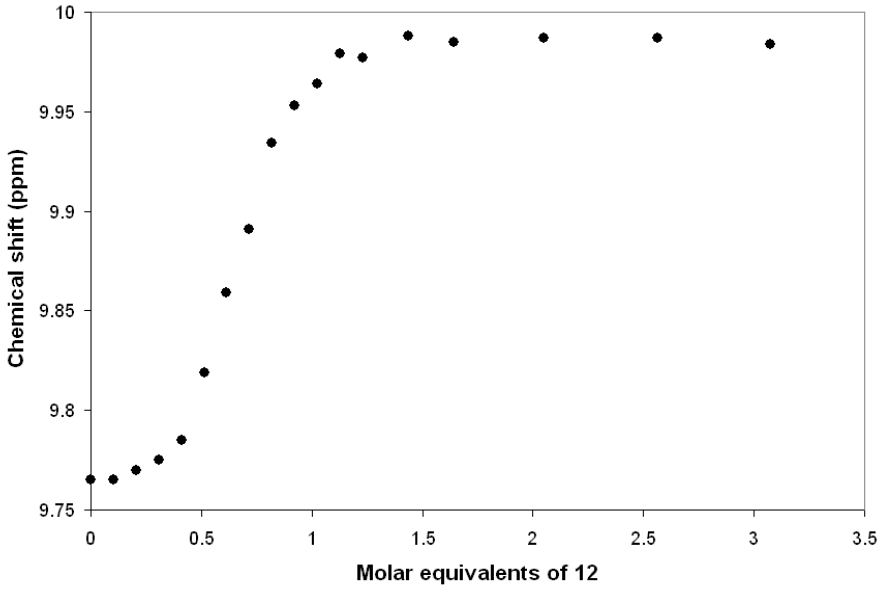


Figure 2.18 | NMR titration (500 MHz, 10% v/v DMSO-d6/$CDCl_3$, 298 K, 2 mM) of 13 with 12, showing a sharp and non-simple isotherm indicative of binding $K_a > 10^5\ M^{-1}$ in addition to multiple equilibria. Proton A of 13 (see Figure 2.4) was followed during the course of the titration.

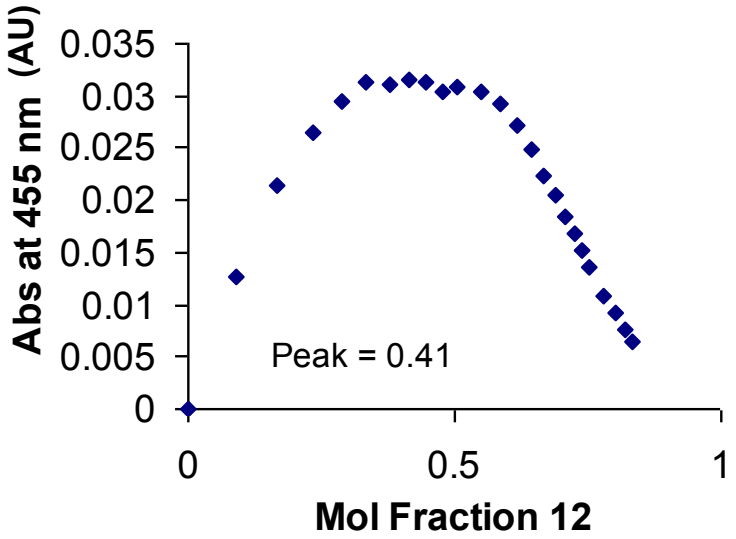


Figure 2.19 | Job plot of 12•13 (10% v/v DMSO/$CHCl_3$, 1×10^{-5} M, UV/vis, 430 nm) showing a binding mode deviating significantly from 1:1.

The unexpected stability of the AAAA-$DDDD^+$ complex **12•13** in very polar solvent systems such as DMSO/$CHCl_3$ and, to a lesser extent, MeCN, may be a result of poor solvation of the uncomplexed components. In particular the high density of interaction

sites in **12**, in either tautomer, may make it difficult to solvate the uncomplexed DDDD^{+} partner effectively. Closely packed solvent heteroatoms around the hydrogen bond donor solvation sites would strongly repel each other. Such ineffective solvation of free hosts or guests is reminiscent of the increase in binding strength found using bulky solvents[33] and in systems with restricted space.[34]

By way of comparison with other quadruple hydrogen bonded systems, Meijer and coworkers have reported studies on the tautomerisation of **9•9** in $CDCl_3$/DMSO-d_6 mixtures.[20] They determined the complex dimerisation constant (K_{dim*}), which is equivalent to the product of the K_{dim} of a single tautomer of **9** and the square of the tautomeric equilibrium constant (K_{taut}) and slightly underestimates K_{dim}, for **9•9** to be ~1.7×10^2 M^{-1} (ΔG = -12.8 kJ mol^{-1}) in 10% DMSO-d_6 in $CDCl_3$ (~50 M^{-1} in neat DMSO-d_6).[32] Zimmerman has reported a K_a of 3.0×10^3 M^{-1} (ΔG = -19.9 kJ mol^{-1}) for an ADDA-DAAD complex in 5% DMSO-d_6 in $CDCl_3$.[22]

2.6 Conclusion

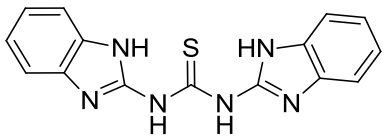
The AAAA-DDDD$^+$ quadruple hydrogen bond complex **12•13** exhibits exceptional complex stability for a small molecule hydrogen bond array in a range of solvent systems ($K_a > 3 \times 10^{12}$ M^{-1} in CH_2Cl_2, 1.5×10^6 M^{-1} in MeCN and 3.4×10^5 M^{-1} in 10% DMSO/$CHCl_3$). Although the modest photostability of **13** may be a concern for some applications, the strength of binding and ease of synthesis of each partner of the AAAA-DDDD$^+$ hydrogen bond motif makes it a promising candidate for incorporation into supramolecular polymers and other functional materials.

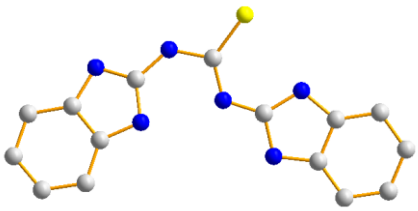
2.7 Experimental procedures and raw data

2.7.1 Synthesis of new compounds

^{1}H and ^{13}C NMR was carried out in $CDCl_3$ at 500/125 or 400/100 MHz. Dry flash chromatography was carried out on silica gel 60 (15-40 µm, Merck) according to the procedure of Harwood.[35] Low resolution mass spectra were obtained with an Agilent Technologies 1200 LC system with positive electrospray ionisation and a 6130 single quadrupole MS detector unless otherwise specified. High resolution mass spectra were obtained with a Bruker 3.0 T Apex II Spectrometer using electrospray ionisation. The following compounds were prepared according to the given literature procedures and were confirmed by ^{1}H NMR: sodium tetrakis[(3,5-trifluoromethyl)phenyl]borate,[36] 1,8-naphthyridine-2,7-diamine.[37,38] All melting points were determined using a Sanyo Gallenkamp apparatus and are uncorrected.

1,3-Bis-(1*H*-benzimidazol-2-yl)-thiourea (25)

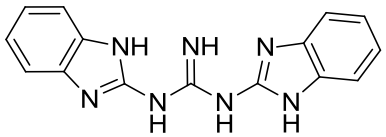




2-Aminobenzimidazole **24** (50 mmol, 6.65 g) and CS_2 (100 mmol, 6.11 mL) were dissolved in pyridine (20 mL) and heated at reflux for 18 h. On cooling, a precipitate formed which was collected and washed with CH_2Cl_2 to give **25** as a colourless crystalline solid (6.28 g, 81%). mp: 272-274 °C (pyridine). ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.8-7.4 (br.s, 4H), 7.4-7.0 (br.s, 4H). MS: 308 (M^+, 10%), 175 (100), 105 (18). *M/z* calcd for $C_{15}H_{12}N_6S$ 308.0839; found 308.0838. Due to rotational broadening, no ^{13}C data could be collected and the atomic connectivity was confirmed by X-ray diffraction (single crystal, slow evaporation of a DMF solution of **25**) by

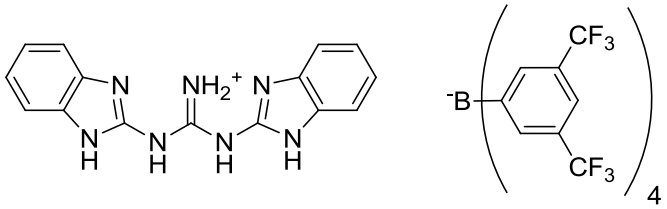
Prof. Alex Slawin (St. Andrews) however experimental details for this have been lost. The compound was also further characterised by conversion to **26** and **12**.

N,*N*'-Bis-(1*H*-benzimidazol-2-yl)-guanidine (26)



The method of Ashworth et al. was adapted.[39] Compound **25** (2.00 g, 6.5 mmol) was suspended in $CHCl_3$ (40 mL) and to this was added HgO (2.00 g, 9.2 mmol) and methanolic NH_3 (2 M, 40 mL). The reaction mixture was stirred at r.t. for 3 h, filtered through celite and concentrated under reduced pressure. The resulting solid was dissolved in acetic acid (2 M, 50 mL) and stirred for 1 h, then filtered through celite and adjusted to pH 8.0 by addition of NaOH (10 M). The precipitate was collected, washed with water and dried. The solid was suspended in Et_2O:MeOH (7.5:1, 200 mL) and then filtered and concentrated under reduced pressure to give **26** as a pale tan solid (944 mg, 50%). mp: 277 °C (Et_2O/MeOH). ^{1}H NMR (500 MHz, DMSO-d_6): δ 12.5-12.0 (br.s, 2H), 9.5-9.0 (br.s, 2H), 7.34 (dd, J = 5.6 Hz, 3.1 Hz, 4H), 7.03 (dd, J = 5.6 Hz, 3.1 Hz, 4H). MS: 291 (M^+, 95%), 133 (89), 78 (89), 63 (100). *M/z* calcd for $C_{15}H_{13}N_7$ 291.1227; found 291.1226.

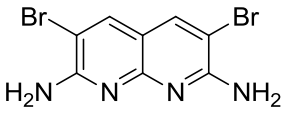
N,*N*'-Bis-(1*H*-benzimidazol-2-yl)-guanidinium tetrakis[(3,5-trifluoro-methyl)phenyl]borate (12)



Compound **26** (300 mg, 1.03 mmol) was dissolved in aqueous acetic acid (8 M, 20 mL) and filtered. To this was added a filtered solution of sodium tetrakis[(3,5-trifluoromethyl)phenyl]borate[36] (1.00 g, 1.13 mmol) dissolved in the minimum

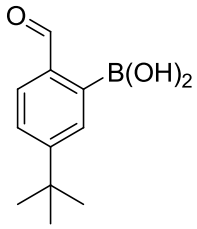
amount of aqueous acetic acid (8 M) and the solution stirred at room temperature until no further precipitate formed. The precipitate was collected on celite and washed with copious water, then CH_2Cl_2. The organic washings were concentrated under reduced pressure to give a foam which was dried *in vacuo* at 40 °C and 1 millibar for 18-72 h, until no traces of acetic acid remained by ^{1}H NMR spectroscopy to give **12** as a colorless solid (552 mg, 40%). mp: 87-91 °C (AcOH/H_2O). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.71 (s, 8H), 7.49 (s, 4H), 7.33 (m, 4H), 7.26 (m, 4H), 5.18 (br.s, 2H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 161.0 (m), 150.6, 134.7, 129.0 (m), 125.6, 125.35, 123.4, 121.3, 117.6, 111.6. MS (APCI -ve) 862 (M^-, 23%), 863 ((M – Ar)$^-$ 100%), 864 (32). MS (APCI +ve): 292 (M^+H, 100%), 149 (30). *M/z* calcd. For cation $C_{15}H_{14}N_7^+$ 292.1305; found 292.1306. *M/z* calcd. for anion $C_{32}H_{12}BF_{24}^-$ 863.0654; found 863.0615.

3,6-Dibromo-1,8-naphthyridine-2,7-diamine (18)



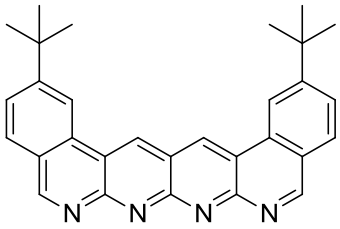
1,8-Naphthyridine-2,7-diamine[37,38] **31** (1.86 g, 11.6 mmol) was dissolved in dry DMF (25 mL). To this was added dropwise over 1 h at -10 °C *N*-bromosuccinimide (4.23 g, 23.8 mmol) in DMF (25 mL). The mixture was then stirred at room temperature for 2 h, and the DMF removed under reduced pressure. The resulting solid was stirred in CH_2Cl_2 (25 mL) for 16 h, and collected by filtration to give **18** as a dark brown solid (3.45 g, 94%). mp: >340 °C (DMF). ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.28 (s, 2H), 7.57 (br.s, 4H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 179.9, 162.8, 156.5, 141.0, 101.3. MS: 316 (53), 318 (M^+, 100), 320 (57), 158 (43), 99 (47). *M/z* calcd. for $C_8H_6{}^{79}Br_2N_4$: 315.8954; found 315.8953.

5-*tert*-Butyl-2-formylboronic acid (33)



The procedure of Comins and Brown was adapted.[40] To a solution of *N,N,N′*-trimethylethylenediamine (17 mmol, 2.16 mL) in dry THF (40 mL) at -30 °C was added *n*-BuLi (16 mmol, 1.6 M in hexanes, 10.0 mL) dropwise with stirring. After 30 min, 4-*tert*-butylbenzaldehyde **20** (15 mmol, 2.50 mL) was added and the mixture stirred for a further 30 min. Further *n*-BuLi (45 mmol, 1.6 M in hexanes, 28 mL) was added and the reaction mixture allowed to warm to room temperature overnight with stirring. The mixture was again cooled to -30 °C, anhydrous $B(OMe)_3$ (90 mmol, 10 mL) was added and the mixture once more allowed to warm to room temperature overnight with stirring. The reaction mixture was poured into cold stirred HCl (100 mL, 2M) and stirred for 30 min, then extracted with diethyl ether (3 × 150 mL), washed with brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure to give the crude product as a dark oil (2.91 g). This was purified by dry flash chromatography (15% acetone/*i*-hexane v/v) to give **33** as a dark solid (1.12 g, 36%). mp: 86-90 °C (acetone/*i*-hexane). ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.87 (s, 1H), 8.35 (d, *J* = 1.9 Hz, 1H), 7.85 (d, *J* = 8.0 Hz, 1H), 7.71 (dd, *J* = 8.0 Hz, *J* = 1.9 Hz, 1H), 7.31 (br.s, 2H), 1.39 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 198.1, 158.2, 139.0, 137.5, 136.1, 127.9, 35.8, 31.0 (*C*-$B(OH)_2$ not observed). MS: 206 (M^+, 31%), 191 (100), 149 (66). *M/z* calcd. For $C_{11}H_{15}BO_3$: 206.1109; found 206.1107.

Hexacene (13)



A solution of **18** (185 mg, 0.58 mmol), **33** (240 mg, 1.16 mmol) in DME (5 mL) and an aqueous solution of Na_2CO_3 (1.5 M, 5 mL) was degassed with nitrogen for 5 min. To this was added $Pd(PPh_3)_4$ (0.06 mmol, 50 mg). The solution was refluxed for 90 min, and then diluted with Na_2CO_3 solution (1.5 M, 20 mL) and extracted with 3 × CH_2Cl_2 (20 mL). The combined extracts were dried over $MgSO_4$, filtered, concentrated under reduced pressure and purified by dry flash chromatography (4% MeOH/$CHCl_3$ v/v) to give **13** as a yellow solid (70 mg, 27%). A sample was purified by trituration with $CHCl_3$ followed by preparative TLC (84.5:15:0.5 EtOAc:MeOH:Et_3N v/v, *R.f.* 0–0.2) to give **13** as a tan solid (25.4 mg, 10%). mp: >340 °C (CH_2Cl_2). ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.33 (s, 2H), 9.68 (s, 2H), 8.97 (d, J = 1.0 Hz, 2H), 8.27 (d, J = 8.2 Hz, 2H), 8.04 (dd, J = 8.2 Hz, 1.0 Hz, 2H), 1.56 (s, 18H). ^{13}C NMR (125 MHz, DMSO-d_6): δ 161.7, 156.0, 155.9, 154.0, 135.0, 131.5, 129.6, 127.2, 123.4, 120.1, 119.0, 118.8, 35.4, 31.0. MS (APCI +ve): 445 (M^+H, 100%). *M/z* calcd for $C_{30}H_{28}N_4$ 444.2308; found 444.2302.

2.7.2 UV/vis titrations to determine binding constants

Titrations to evaluate the strength of binding were carried out by UV/vis spectroscopy. NMR titrations, though the workhorse of binding constant evaluations, have a number of limitations that made them particularly unsuitable for our purposes. The higher sensitivity of UV/vis meant that experiments could be carried out over a wider concentration range, and with any non-absorbing solvent. Additionally, acquisition times are longer and less data is returned overall for NMR experiments. Finally, the NMR spectrum of **12** and **13** were unobtainable in CD_3CN.

UV/vis experiments were performed on a Varian Cary Bio-50 UV/vis spectrometer at 298K, surveying a range of 250-550 nm at at a resolution of 1.0 nm and an integration time of 0.1 s. Quartz cuvettes with Teflon stoppers were used; unless otherwise specified the cell path length was 10.0 mm and the working volume was 3.5 mL. Measurements of volumes > 250 μL were carried out using Hamilton gastight syringes of 1, 2.5, or 10 mL volume, and measurements of volumes < 250 μL were carried out using Hamilton microliter syringes of 250, 100, 50, 25, or 10 μL volume, in each case using the smallest syringe that would contain the entire measurement volume in order to minimise error (for example, 2.0 mL of stock solvent would be measured with a 2.5 mL Hamilton gastight syringe, and a titration aliquot of 15 μL would be measured with a 25 μL Hamilton microliter syringe.)

A "host" and a "guest" were designated for each titration experiment. The hydrogen bond donors (**12** or **7**) were called guests, and the hydrogen bond acceptors (**6** or **13**) were called hosts. In a host-guest titration, a measurable change in the host is followed which is dependent on the guest. The host should be measured without interference from the guest; in these experiments, **6** and **13** (hosts) absorbed in the UV region > 380 nm, whereas **12** and **7** (guests) only absorbed < 380 nm. When performing data-fitting, only data collected above 380 nm was used in order to specifically track the host.

For each individual titration, the experiment was carried out as follows: A host molecule was dissolved in the chosen solvent for that experiment, and diluted to 10 mL of a specified concentration (given for each set of data). Then, a portion of that host was used to dissolve a quantity of guest such that the final concentration of the guest was between ten and twenty times that of the host. In this way, whenever a solution of guest was added to the host, the concentration of host did not change – the guest contains a "background concentration" of host.

During a titration, a volume of guest solution was measured, and added to the UV/vis cell containing the host. The cell was then stoppered and inverted several times to ensure complete mixing. In several cases, individual scans in the middle of a titration were re-run after a wait of several minutes to yield identical spectra, showing that equilibration is fast and that no other decay processes take place on the same timescale as the titrations. During a titration, the total number of equivalents of guest added (relative to host) were as follows unless otherwise specified: 0, 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 2, 2.5, 3, 4, 6. Care was taken to note that, due to the background concentration of host, a given volume of guest solution did not always correspond to the same number of equivalents of host. Each titration was carried out in triplicate, and values of the binding constant agreed to within 15%.

The exact volumes of guest added during each titration were noted and used to calculate a list of concentrations of host and guest, which is the format the data modelling software took. As long as the correct concentrations of host and guest were assigned to the individual UV/vis spectral data, the modelling software was able to calculate an accurate binding constant. Spectral analyses were performed using the ReactLab^{TM} Equilibria analysis suite (Jplus Consulting, www.jplusconsulting.com).

2.7.3 Evaluating binding constants in CH_2Cl_2

In this competition experiment, a 1:1 mixture of **6** and **13** was used as the host, and a solution of **7** was used as the guest. The fitting software was provided with the previously-established value for the binding constant of **6•7**, and returned values of the binding constant of **6•13** without any need to manually de-convolute the data. The fitting software returns values of $Log_{10}(K)$, and these are given plus or minus one standard deviation, which was within 1% in each case (Table 1).

Equilibrium modelled	Concentration	$Log_{10}(K)$ for 3 runs			Average $log_{10}(K)$
		1	2	3	
6•7 + **13** ↔ **6·13** + **7**	57 μM	9.96	10.00	10.09	10.02 ± 0.065

Table 1 \| Values of binding constant for a competition between 6•7 and 13, evaluated in CH_2Cl_2 at 298 K with a path length of 10.0 mm, at the specified constant concentration of 6•7.

2.7.4 Evaluating binding constants in MeCN

Four different complexes were investigated: **7•6**, **12•6**, **7•13** and **12•13**. For each complex, experiments were performed in triplicate (Table 2). The fitting software returns values of $Log_{10}(K)$, and these are given plus or minus one standard deviation, which was within 1% in each case. The concentration of host (**6** or **13**) is given, which was kept constant throughout the titration.

Equilibrium modelled	Concentration	$Log_{10}(K)$ for 3 runs			Average $log_{10}(K)$
		1	2	3	
6 + **7** ↔ **6·7**	79 μM	3.78	3.77	3.64	3.73 ± 0.076
6 + **12** ↔ **6·12**	104 μM	4.15	4.12	4.19	4.15 ± 0.030
13 + **7** ↔ **13·7**	6.4 μM	4.79	4.90	4.74	4.81 ± 0.081
13 + **12** ↔ **13·12**	3.6 μM	6.21	6.28	6.03	6.17 ± 0.122
13·12 + **13** ↔**(13·12)·13**[a]	3.6 μM	5.52	5.56	5.53	5.53 ± 0.032

Table 2 | Values of binding constant for every permutation, matched and mismatched, where 6 or 13 were the acceptor, and 7 or 12 were the acceptor, evaluated in MeCN at 298 K with a path length of 10.0 mm, at the specified constant concentration of 6 or 13. (a) values of $Log_{10}(K)$ are those determined from the same dataset as the corresponding 1:1 equilibrium. Entirely separate acquisition and analysis was carried out in triplicate, in good agreement.

The stoichiometry of each equilibria was confirmed by Job Plots, which were calculated from the existing data by the method of MacCarthy.[41, 42]

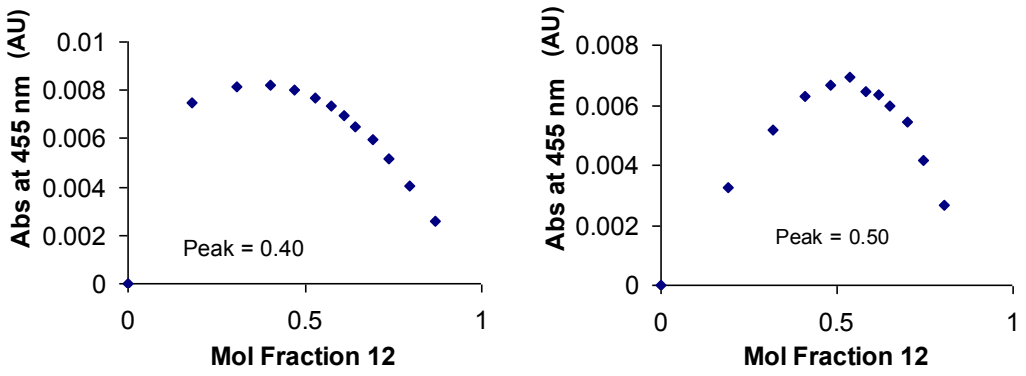


Figure 2.20 | Job plot of 12•13 (MeCN, 5 × 10^{-5} M, UV/vis, 460 nm) showing a mixed binding mode (left). Job plot of 12•6 (MeCN, 7 × 10^{-4} M, UV/vis, 370 nm) showing a 1:1 binding mode (right).

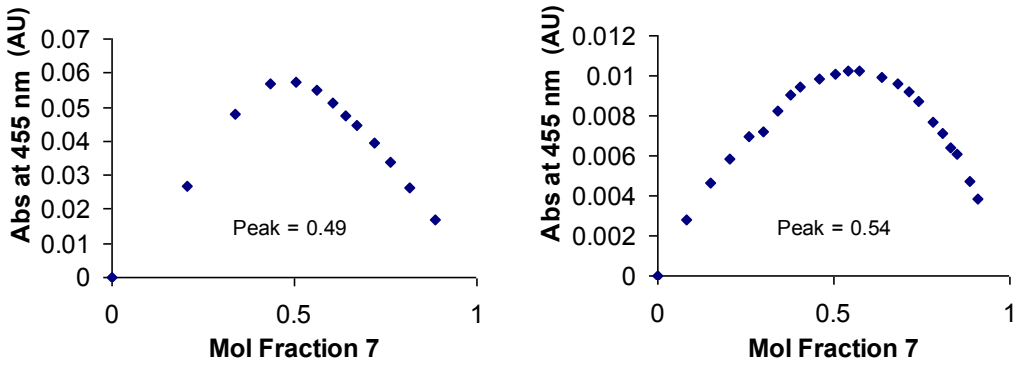


Figure 2.21 | Job plot of 7•6 (MeCN, 8 × 10^{-4} M, UV/vis, 370 nm) showing a 1:1 binding mode (left). Job plot of 7•13 (MeCN, 5 × 10^{-5} M, UV/vis, 450 nm) showing a 1:1 binding mode (right).

The stoichiometry of each equilibria was also confirmed as follows: For each experiment, identical fitting procedures were carried out with and without a programmed model of a 2:1 equilibrium of Acceptor to Donor. For **7•6**, **12•6**, and **7•13**, the fits for a 2:1 equilibrium either did not converge or converged to give the binding constant of the 2:1 equilibrium as 0. For **12·13**, the goodness-of-fit parameter was almost 10 times worse than when a 2:1 equilibrium was omitted. Additionally, the peak residuals were non-random, and typically 3-5 times larger for any given wavelength, with some over- or underestimating the absorbance by >5% (example given in Figure 2.22).

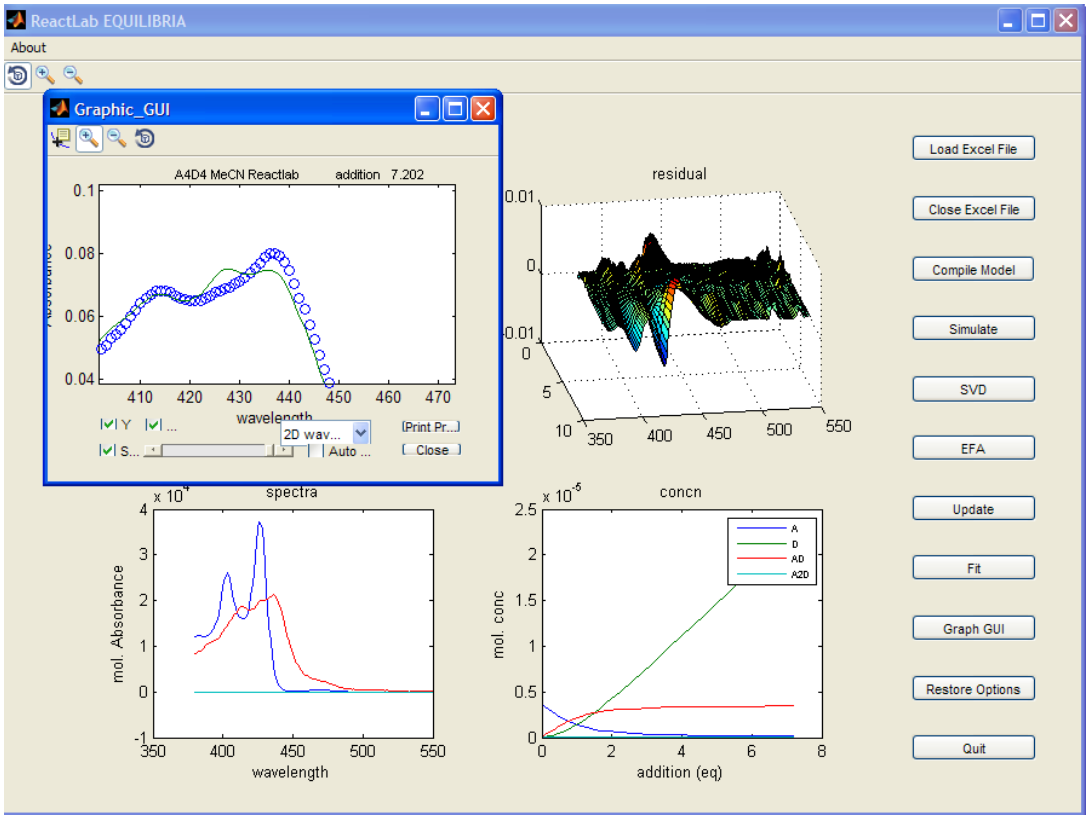


Figure 2.22 | UV/vis titration of 13 with 12 in MeCN, omitting the presence of a 2:1 binding mode, ReactLab Working Window. The fit is considerably worse as evidenced by the 3D plot of residuals (top-right), which shows non-random behaviour. Top left, a detail from the titration showing a poor fit of the peaks at 425 and 440 nm (Green line = predicted spectrum, blue circles = data).

2.7.5 Evaluating binding constants in 10% DMSO/$CHCl_3$

For complex **12•13** in 10% DMSO/$CHCl_3$, titrations were performed in triplicate (Table 3). The fitting software returns values of $Log_{10}(K)$, and these are given plus or minus one standard deviation, which was within 1% in each case. The concentration **13** is given, which was kept constant throughout the titration.

Equilibrium modelled	Concentration	$Log_{10}(K)$ for 3 runs			Average $log_{10}(K)$
		1	2	3	
13 + **12** ↔ **13·12**	10 μM	5.51	5.51	5.56	5.53 ± 0.031
13·12 + **13** ↔**(13·12)·13**[a]	10 μM	5.13	5.14	5.13	5.13 ± 0.003

Table 3 \| Values of binding constant 13 with 12 in 10% DMSO/$CHCl_3$ by UV/vis. (a) Values of $Log_{10}(K)$ are those determined from the same dataset the corresponding 1:1 equilibrium. Entirely separate acquisition and analysis was carried out in triplicate, in good agreement.

2.7.6 Crystal data and structure refinement for 12•13

Identification code	[5.4]	
Empirical formula	C77.25 H55 B F24 N11 O0.25	
Formula weight	1608.13	
Temperature	173(2) K	
Wavelength	1.54178 Å	
Crystal system	Monoclinic	
Space group	P2(1)/n	
Unit cell dimensions	a = 9.156(3) Å	α = 90°.
	b = 32.885(11) Å	β = 92.749(9)°.
	c = 27.081(9) Å	γ = 90°.
Volume	8145(5) $Å^3$	
Z	4	
Density (calculated)	1.311 Mg/m^3	
Absorption coefficient	1.029 mm^{-1}	
F(000)	3274	
Crystal size	0.20 x 0.10 x 0.01 mm^3	
Theta range for data collection	2.69 to 68.69°.	
Index ranges	-11<=h<=11, -39<=k<=38, -32<=l<=32	
Reflections collected	107989	
Independent reflections	14892 [R(int) = 0.2520]	
Completeness to theta = 67.00°	99.8 %	
Absorption correction	Multiscan	
Max. and min. transmission	1.000 and 0.679	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	14892 / 252 / 1062	
Goodness-of-fit on F2	1.584	
Final R indices [I>2sigma(I)]	R1 = 0.2181, wR2 = 0.4941	
R indices (all data)	R1 = 0.2981, wR2 = 0.5297	
Largest diff. peak and hole	1.074 and -0.846 e.$Å^{-3}$	

Hydrogen bond distances and angles [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(1)-H(1A)...N(21)	0.98	1.81	2.768(11)	165.7
N(2)-H(2A)...N(22)	0.98	2.00	2.979(11)	176.9
N(8)-H(8A)...N(16)	0.98	1.84	2.601(11)	132.5
N(8)-H(8B)...N(3)	0.98	1.88	2.630(11)	131.5
N(8)-H(8A)...F(9)#1	0.98	2.55	3.192(12)	123.0
N(8)-H(8B)...F(7)#1	0.98	2.22	2.840(11)	119.7
N(9)-H(9A)...N(23)	0.98	2.03	3.006(11)	173.8
N(11)-H(11A)...N(24)	0.98	1.77	2.745(11)	173.9

Symmetry transformations used to generate equivalent atoms:

#1 x-1/2,-y+1/2,z-1/2

2.8 References

(1) de Greef, T. F. A. et al. *Chem. Rev.* **109**, 5697–5754 (2009).

(2) Fathalla, M., Lawrence, C. M., Zhang, N., Sessler, J. L. & Jayawickramarajah, J. *Chem. Soc. Rev.* **38**, 1608–1620 (2009).

(3) de Greef, T. F. A. & Meijer, E. W. *Nature* **453**, 171–173 (2008).

(4) Wilson, A. J. *Soft Matter* **3**, 409–425 (2007).

(5) Zimmerman, S. C. & Corbin, P. S. *Struct. Bonding (Berlin)* **96**, 63–94 (2000).

(6) Jorgensen, W. L. & Pranata, J. *J. Am. Chem. Soc.* **112**, 2008–2010 (1990).

(7) Popelier, P. L. A. & Joubert, L. *J. Am. Chem. Soc*. **124**, 8725–8729 (2002).

(8) Quinn, J. R., Zimmerman, S. C., Del Bene, J. E. & Shavitt, I. *J. Am. Chem. Soc.* **129**, 934–941 (2007).

(9) Kyogoku, Y., Lord, R. C. & Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **57**, 250–257 (1967).

(10) Kyogoku, Y., Lord, R.s C. & Rich, A. *Biochim. Biophys. Acta*. **179**, 10–17 (1969).

(11) Murray, T. J &. Zimmerman, S. C. *J. Am. Chem. Soc*. **114**, 4010–4011 (1992).

(12) Bell, D. A. & Anslyn, E. V. *Tetrahedon* **51**, 7161–7172 (1995).

(13) Djurdjevic, S. et al. *J. Am. Chem. Soc.* **129**, 476–477 (2007).

(14) Blight, B. A. et al. *J. Am Chem Soc*. **131**, 14116–14122 (2009).

(15) Sartorius, J. & Schneider, H.-J. *Chem. Eur. J.* **2**, 1446–1452 (1996).

(16) Beijer, F. H., Kooijman, H., Spek, A. L., Sijbesma, R. P. & Meijer, E. W. *Angew. Chem. Int. Ed.* **37**, 75–78 (1998).

(17) Sijbesma, R. P. et al. *Science* **278**, 1601–1604 (1997).

(18) Ligthart, G. B. W. L., Ohkawa, H., Sijbesma, R. P. & Meijer, E. W. *J. Am. Chem. Soc*. **127**, 810–811 (2005).

(19) Lafitte, V. G. H., et al. *J. Am. Chem. Soc.* **128**, 6544–6545 (2006).

(20) Beijer, F. H., Sijbesma, R. P., Kooijman, H., Spek, A. L. & Meijer, E. W. *J. Am. Chem. Soc*. **120**, 6761–6769 (1998).

(21) Greco, E. et al. *New J. Chem*. DOI: 10.1039/C0NJ00197J.

(22) Corbin, P. S. & Zimmerman, S. C. *J. Am. Chem. Soc.* **120**, 9710–9711 (1998).

(23) Park, T., Zimmerman, S. C. & Nakashima, S. *J. Am. Chem. Soc*. **127**, 6520–6521 (2005).

(24) Kuykendall, D. W., Anderson, C. A. & Zimmerman, S. C. *Org. Lett.* **11**, 61–64 (2009).

(25) Brammer, S., Lüning, U. & Kühl, C. *Eur. J. Org. Chem*. 4054–4062 (2002).

(26) Taubitz, J. & Lüning, U. *Aust. J. Chem.* **62**, 1550–1555 (2009).

(27) Hisamatsu, Y., Shirai, N., Ikeda, S.-i. & Odashima, K. *Org. Lett.* **12**, 1776–1779 (2010).

(28) McGhee, A. M.; Kilner, C.; Wilson, A. J. *Chem. Commun.* 344-346 (2008).

(29) Connors, K. A. Binding Constants: The Measurement of Molecular Complex Stability (Wiley-Interscience, New York, 1987)

(30) Hunter, C. A. *Angew. Chem., Int. Ed*. **43**, 5310–5324 (2004).

(31) Hunter, C. A., Hutchinson, J. J., Varley, L. M. manuscript in preparation.

(32) Lafitte, V. G. H.; Aliev, A. E.; Hailes, H. C.; Bala, K.; Golding, P. *J. Org. Chem*. **70**, 2701-2707 (2005).

(33) Chapman, K. T. & Still, W. C. *J. Am. Chem. Soc*. **111**, 3075–3077 (1989).

(34) Mecozzi, S. & Rebek, Jr., J. *Chem. Eur. J.* **4**, 1016–1022 (1998).

(35) Harwood, L. M. *Aldrichimica Acta* **18**, 25 (1985).

(36) Yakelis, N. A. & Bergman, R. G. *Organometallics* **24**, 3579-3581, (2005).

(37) Park, T., Mayer, M. F., Nakashima, S. & Zimmerman, S. C. *Synlett*, 1435-1436, (2005).

(38) A more efficient synthesis has since been published: Goswami, S., Mukherjee, R., Jana, S., Maity, A. C. & Adak, A. K. *Molecules* **10**, 929-936 (2005).

(39) Ashworth, R. D., Crowther, A. F., Curd, F. H. S. & Rose, F. L. *J. Chem. Soc.*, 581-586 (1948).

(40) Comins, D. L. & Brown, J. D. *Tetrahedron Lett.* **24**, 5465-5468 (1983).

(41) Hill, Z. D. & MacCarthy, P. *J. Chem. Ed.* **63**, 162-167, (1986).

(42) MacCarthy, P. *J. Anal. Chem.*, **50**, 2165, (1978)

Chapter III

AAAA-DDDD Quadruple Hydrogen Bond Arrays Featuring NH•••N and CH•••N Hydrogen Bonds

Portions of this chapter will be published as: AAAA-DDDD quadruple hydrogen bond arrays featuring NH•••N and CH•••N hydrogen bonds, Leigh, D. A.; Robertson, C. C.; Slawin, A. M. Z.; Thomson, P. I. T. *Journal of the American Chemical Society,* submitted.

Acknowledgements

The following people are gratefully acknowledged for their contribution to this chapter: Professors Hamish McNab and David Leigh proposed the research, contributed to the writing of the paper and provided many useful discussions. Dr Craig Robertson contributed to discussion and preparation of the original paper including figures 3.11 and 3.14.

3.1 Synopsis

This chapter concerns the synthesis and subsequent analysis of further contiguous quadruple hydrogen bonded arrays, continuing from the work of chapter II.

Previously, we reported an extremely strong AAAA-DDDD quadruple NH•••N hydrogen bond array ***5•4*** *with a* $K_a = 1.5 \times 10^6\ M^{-1}$ *in* CH_3CN ($K_a > 3 \times 10^{12}\ M^{-1}$ *in* CH_2Cl_2*). Herein, we report that changing the two benzimidazole groups of the DDDD unit to triazole groups will replace two of the NH•••N hydrogen bonds with CH•••N interactions (complex* ***5•6****), but only reduces the association constant in* CH_3CN *by two orders of magnitude* ($K_a = 2.6 \times 10^4\ M^{-1}$ *in* CH_3CN; $K_a > 1 \times 10^7\ M^{-1}$ *in* CH_2Cl_2*). Related complexes without the triazole groups range in* K_a *from 18-270* M^{-1} *in* CH_3CN*, suggesting that the CH•••N interactions can be considered part of an AAAA-DDDD quadruple hydrogen bonding array. The NH•••N/CH•••N AAAA-DDDD motif can be repeatedly switched 'on' and 'off' in* $CDCl_3$ *through successive additions of acid and base.*

3.2 Introduction

Whilst individual hydrogen bonds are generally weak with short lifetimes, their enthalpy of formation is additive, meaning that multipoint hydrogen bonding arrays[1-8] can effectively hold together supramolecular assemblies and dynamic materials.[9-14] As previously discussed (Chapter I), secondary electrostatic interactions between adjacent hydrogen bonds have a significant effect on the stability of a supramolecular complex.[15,16] The binding strength is theoretically maximised if all the hydrogen-bond donors (D) are on one component and all the hydrogen-bond acceptors (A) are on the other,[15,16] a trend that has been experimentally demonstrated with triple[19-21] and quadruple[18] hydrogen bond motifs involving N-H donors and N or O hydrogen bond acceptors (e.g. AADD-DDAA **1•1** < ADDA-DAAD **2•3** < AAAA-DDDD **5•4**, Figure 3.1).

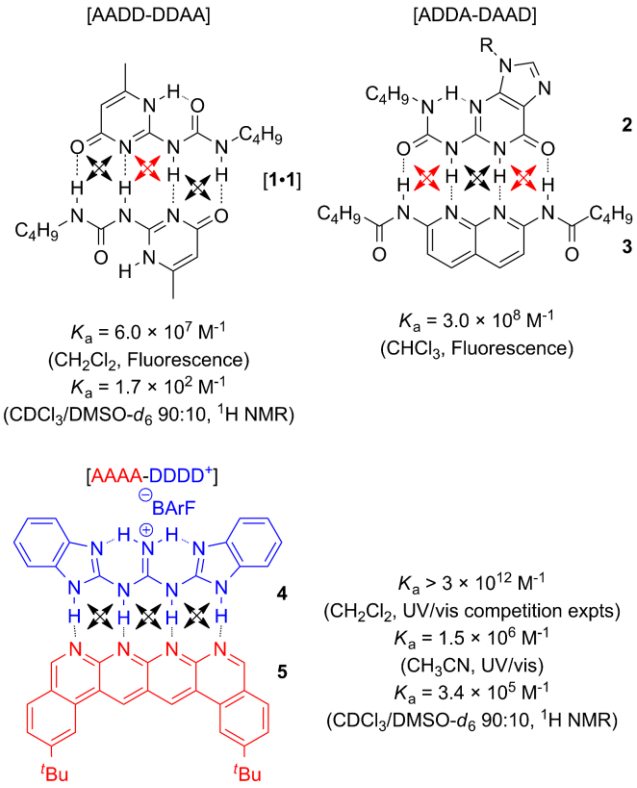


Figure 3.1 | Examples of quadruple hydrogen bond arrays and their association constants in various solvents (method of binding constant determination shown in parentheses): AADD-DDAA 1•1[2], ADDA-DAAD 2•3[17] and AAAA-DDDD 5•4[18]. Black arrows show stabilizing secondary electrostatic interactions between adjacent hydrogen bonding sites; red arrows show destabilizing secondary interactions.

Although considered to be generally weaker, C-H•••N/O hydrogen bonds have recently been shown to play important roles in molecular recognition and assembly processes,[32] including protein-protein interactions,[46] anion recognition[36-39] and extended crystal lattices.[33]

We recently reported an extremely strong (K_a > 3 × 10^{12} M^{-1} in CH_2Cl_2) AAAA-DDDD quadruple hydrogen bonding array (**5•4**, Figure 3.1) based on four NH•••N hydrogen bonds.[18] Here we report the effect on the strength of binding of the quadruple hydrogen bond array of replacing two of the NH•••N hydrogen bonds for CH•••N interactions in **6•5**.

3.3 Design and Synthesis of arrays bearing CH•••N Hydrogen bonds

3.3.1 Design of Hydrogen Bond Donor 6

The N-H of the benzimidazole groups of DDDD template **4** still partake in short near-linear hydrogen bonds as evidenced in the crystal structure of the complex (Chapter II, Section 2.4.3), despite being in five-membered rings which are not optimal for presenting the NH groups parallel to each other. The key features of this array were the guanidinium core, complexed by a pair of hydrogen bond acceptors. With these key features intact, the peripheral hydrogen bonds of any recognition array can take almost any form and still be held in close proximity by the strong hydrogen bonding in the centre of the array. In initial work, several structures were proposed (Figure 3.2).

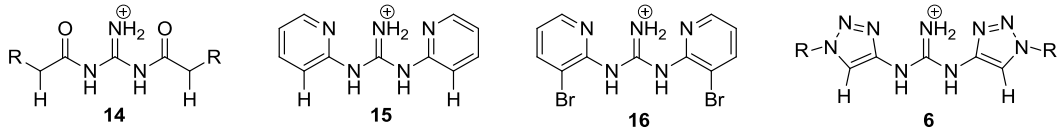
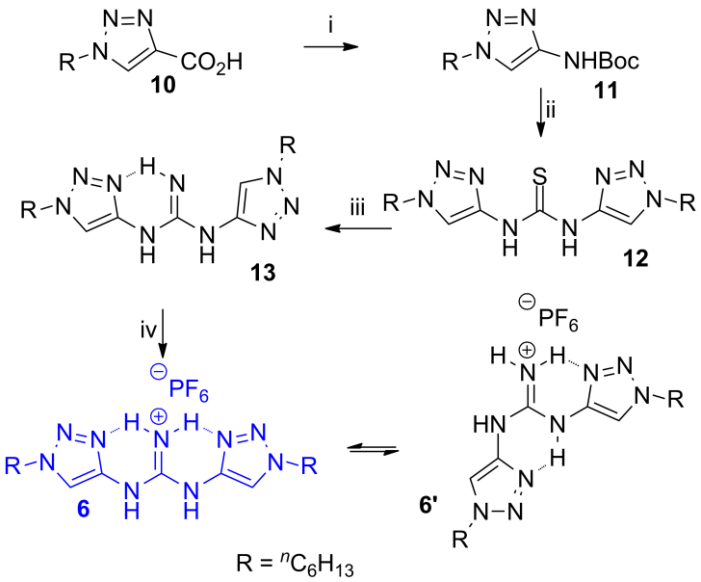


Figure 3.2 | Possible hydrogen or halogen-bonding motifs that would present a "non-conventional" quadruple hydrogen bonding array.

Compound **14** would hydrogen bond through the acidic α-protons, but was thought to suffer from non-ideal sterics. Compound **15** would lie planar, but would not be directly comparable to other five-membered ring systems. Compound **16** was proposed to study mixed halogen-hydrogen bonding, but could never be successfully synthesised. Triazole **6** most closely mirrored the geometry of the previously-studied benzimidazole, allowing the effects of CH•••N hydrogen bonding to be studied independently of steric considerations, and so it was selected as a primary candidate for synthesis and study, as well as **14**, **15**, and **16**.

3.3.2 Synthesis and Characterisation of Hydrogen Bond Donors

A potential N-H/C-H hydrogen bond donor **6** was synthesised in five steps from known[6] triazole **10** as shown in Scheme 3.1. Amine **11** was produced *via* a Curtius rearrangement followed by deprotection with trifluoroacetic acid and subsequent thiourea formation with carbon disulfide in pyridine (Scheme 3.1, i-ii). Thiourea **12** was transformed into guanidine **13** *via* a carbodiimide intermediate and precipitated as the hexafluorophosphate salt **6** (Scheme 3.1, iii-iv). In principle **6** can exist in another conformer stabilised by intramolecular hydrogen bonding, and this may complicate an NMR spectrum or binding equilibrium (**6** and **6′**, Scheme 3.1).



Scheme 3.1 | Synthesis of N-H/C-H hydrogen bond donor 6. Reagents and conditions: (i) $(PhO)_2PON_3$, tBuOH, Et_3N, 16 h, 100 °C, 33%. (ii) CF_3CO_2H, CH_2Cl_2 then CS_2, pyridine, 16 h, 130 °C, 83% over two steps. (iii) HgO, NH_3, $CHCl_3$/MeOH, 1 h, rt, 91%. (iv) AcOH, $NaPF_6$, 30 min, rt, 70%.

Compounds **14** and **15** was also synthesised by an analogous route, as **14**•PF_6^- and **15**•$BArF^-$ (see Experimental, Section 3.7.1).* Compound **16** could never be successfully synthesised, and was halted after suggestions that the halogen bonding would be ineffective.

* Binding constants of **14**•**5** and **15**•**5** were found to be 5×10^3 M^{-1} and 1.5×10^3 M^{-1} (MeCN, 298 K, UV/vis) respectively (as per Chapter II, Section 2.7.2) but only compound **6** was taken on.

3.4 Evidence of formation of arrays bearing CH•••N Hydrogen bonds

3.4.1 Evidence of formation of 5•6 by NMR

Addition of **6** to **5** in CD_2Cl_2 to form a 1:1 solution immediately led to considerable shifts in the ^1H NMR spectra of both components (Figure 3.3). The triazole C-H protons of **6** (H_a) are broadened and shifted downfield by ~2 ppm, consistent with their very close positioning to, and polarisation by, a region of high electron density (e.g. a heteroatom lone pair). The protons of the pyridyl rings of **5** (H_A and H_B) also undergo substantial shielding and deshielding in the complex. Proton A could be assigned by the presence of a ROESY crosspeak with b (Figure 3.5).

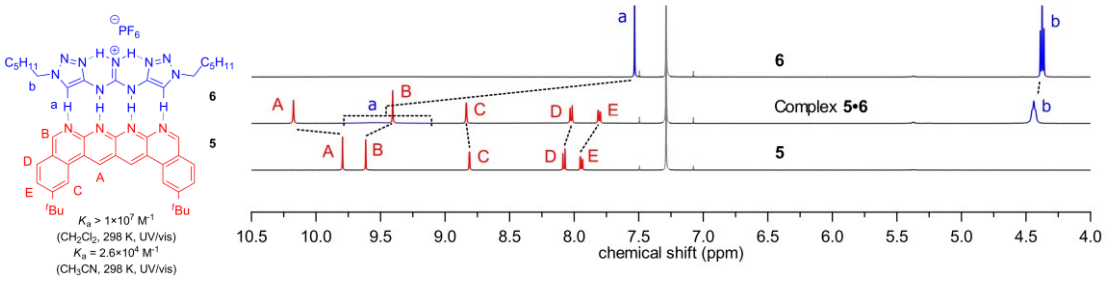


Figure 3.3 | Partial ^1H NMR spectra (500 MHz, CD_2Cl_2, 298 K) of 6 (top), complex 5•6 (middle) and 5 (bottom). Dashed lines show shifts upon formation of complex 5•6. Residual $CHCl_3$ shown in gray.

An intermolecular NOESY correlation between a 1:1 mixture of **6** and **5** was observed in $CDCl_3$ (Figure 3.4). Other intramolecular correlations helped assign the structure of **5** and **6** when complexed, showing the expected correlation between the alkyl chain proton **6b** and the proximal heteroaromatic proton **5B**.

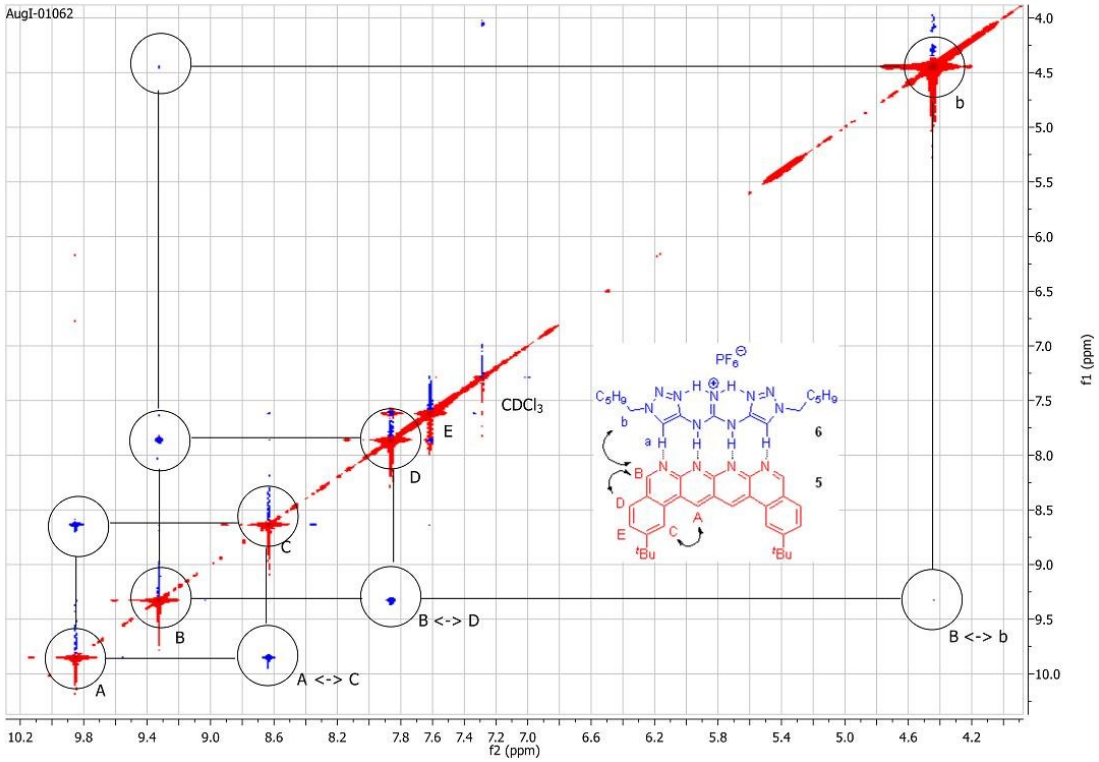


Figure 3.4 | ^{1}H NOESY experiment of 5•6 (5 mM, $CDCl_3$, 298 K, 500 MHz)

In order to observe the behaviour of triazole proton **6a** when complexed, a ROESY experiment was carried out which showed a weak correlation between **6a** and **6b-6c**, placing **6a** at 8.3-8.6 ppm (Figure 3.5). In the 1D ^{1}H spectrum at 298K, the resonance spans over 1 ppm, probably as a result of interconversion between tautomers/conformers that occurs on the NMR timescale (Figure 3.6).

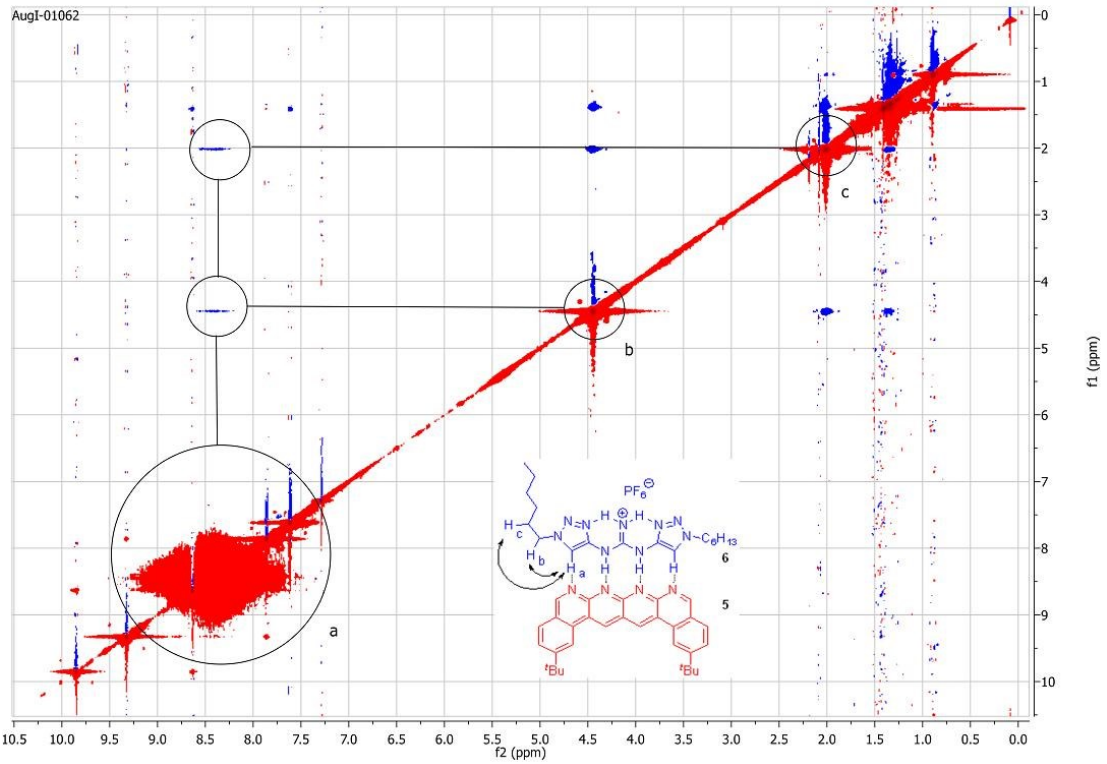


Figure 3.5 | ^{1}H ROESY experiment showing the interaction of 5 and 6 (5 mM, 298 K, 500 MHz)

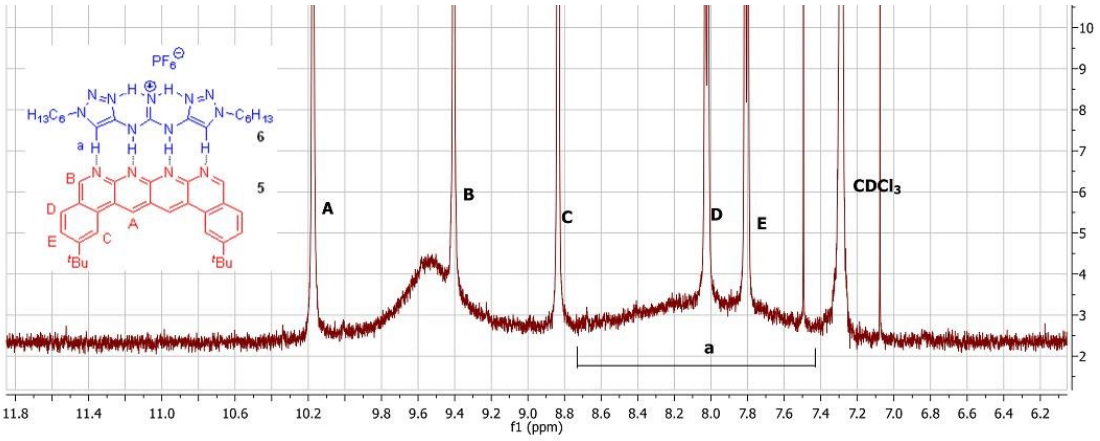


Figure 3.6 | ROSEY and partial 1D ^{1}H NMR spectrum of 5•6 (500 MHz, $CDCl_3$, 298 K). This is the same data as figure 3.3, middle trace, but amplified to show more clearly the broadened baseline resonances. The peak at 9.2 - 9.8 ppm could not be assigned but most likely belongs to one or more of the guanidinium protons.

3.4.2 Evidence of formation of 5•6 by Mass Spectrometry

Complex $[\mathbf{5•6}]^+$ in MeCN (1 mM) was observed by ESI^+ (Thermo Scientific, LCQ Fleet) at *m/z* = 806, as well as free $[\mathbf{6}+H]^+$ at *m/z* = 362 (Figure 3.7). Detailed examinations observed isotopic distribution patterns consistent with those calculated for free $[\mathbf{6}+H]^+$ (Figure 3.8) and $[\mathbf{5•6}]^+$ (Figure 3.9), to a high degree of accuracy.

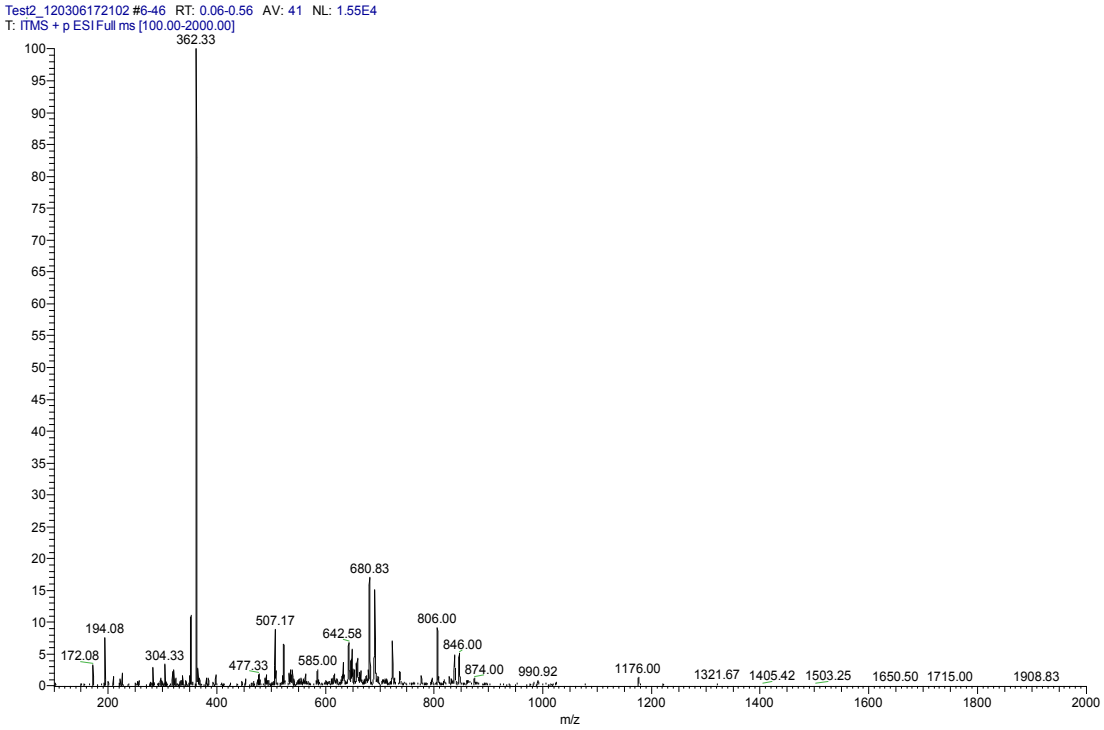


Figure 3.7 | Full low-resolution mass spectrum (ESI+) of complex 5•6 showing $[6+H]^+$ (362) and $[5•6]^+$ (806).

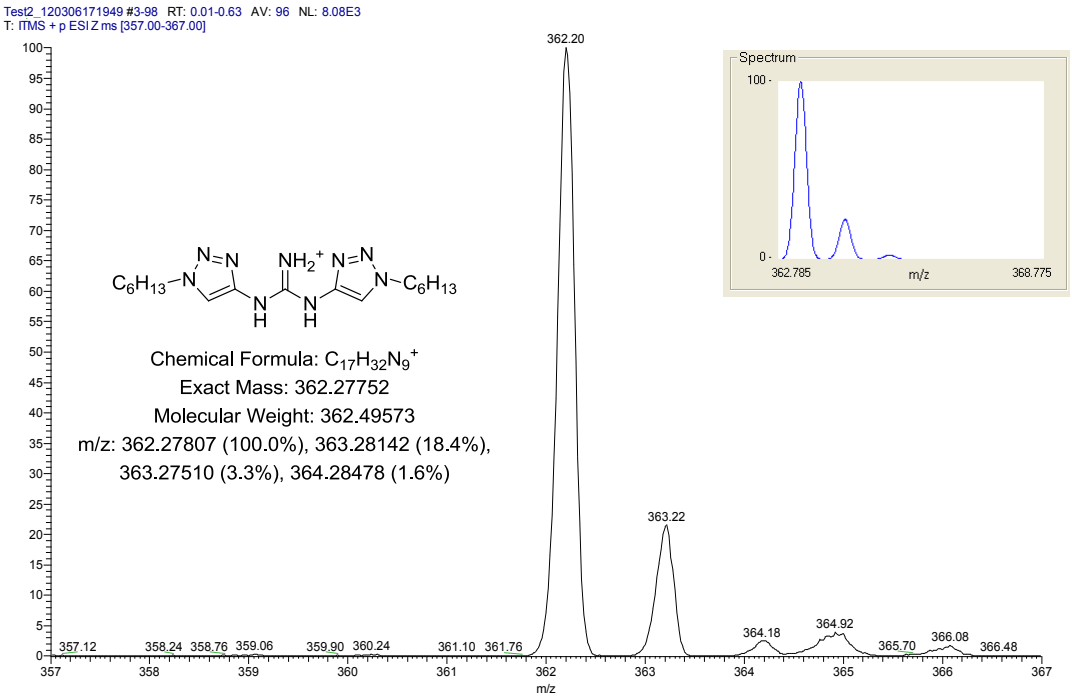


Figure 3.8 | Peak detail of [6+H]$^+$. Inset: structure (left) and predicted isotopic distribution pattern[45] (right).

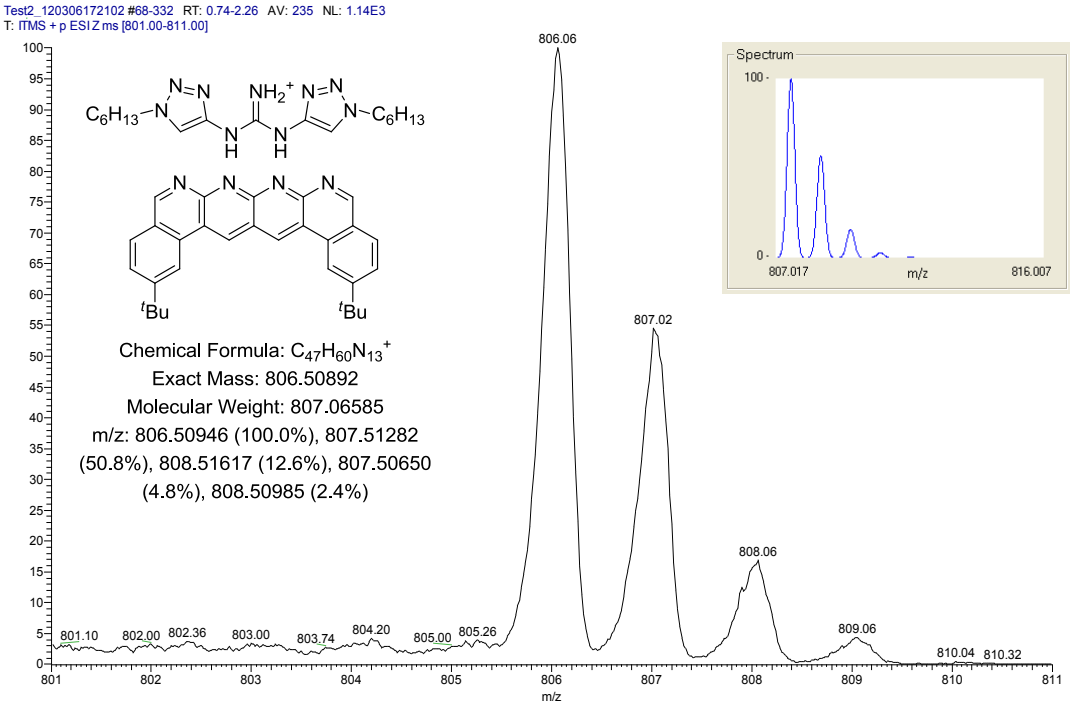


Figure 3.9 | Peak detail of [5•6]$^+$. Inset: structure (left) and predicted isotopic distribution pattern[45] (right).

3.5 Evaluation of binding of arrays bearing CH•••N Hydrogen bonds

3.5.1 Evaluation of 5•6 in MeCN

The association constant of **5•6** was found to be too great to be measured directly in CD_2Cl_2 by ^1H NMR or UV/vis spectroscopy, using the same attempted methods as previously described (Chapter II, Section 2.7.2).

However, a titration of **6** with **5** at 10^{-4} M in CH_3CN (298 K) showed (Figure 3.10) a decrease in UV/vis absorption spectrum of **5** (λ_{max} = 426 nm) accompanied by a new species with a bathochromic shift ($\Delta\lambda$) of 11 nm. From this data the K_a of **5•6** in CH_3CN was determined to be 2.6×10^4 M^{-1} (see section 3.7.2 for details). This means that replacing two of the NH•••N hydrogen bonds in **5•4** ($K_a = 1.5 \times 10^6$ M^{-1}) with the two CH•••N interactions in **5•6** ($K_a = 2.6 \times 10^4$ M^{-1}) results in a mere ~60-fold decrease in the K_a value in CH_3CN and a $\Delta\Delta G$ of 10.2 $kJmol^{-1}$.

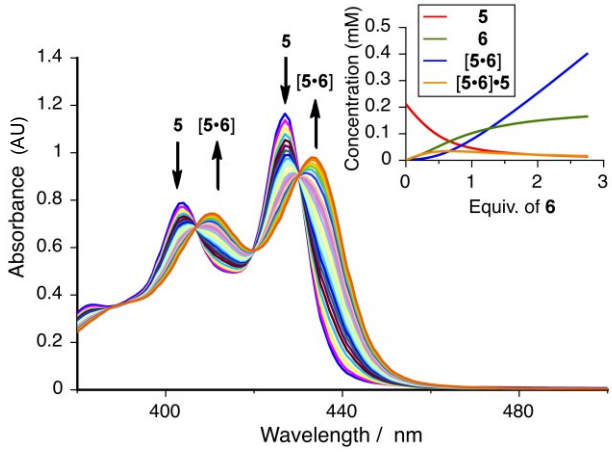


Figure 3.10 | UV/vis titration of 5 with 6 in CH_3CN. UV/vis spectra (2.1×10^{-4} M) of 5 on addition of 6 (0-3 equiv), maintaining the concentration of 5 constant. Changes in absorbance reflect changes in the amount of 5 and 5•6 present during the titration experiment and differences in their UV/vis absorption. Inset: profile of component stoichiometry from fitting software.

3.5.2 Complexes featuring multiple CH•••N and CH•••N interactions

To further probe the contribution of the CH•••N interactions to the overall stability of the hydrogen bonding array, a series of association constants were measured (Figure 3.11) using binding partners with increasing numbers of contiguous acceptor sites (**8** with two acceptor sites and **7** with three) with donors containing two NHs from a central guanidinium core plus either two further NHs (**4**), CHs (**6**) or no additional groups (**9**) able to interact with pyridine sites on their binding partner. The rationale is that with each of **5**, **7** and **8** two pyridine sites would satisfy the hydrogen bond requirements of the guanidinium NH groups of **6**, **4** and **9**, but that **8** has no further pyridine groups to engage in CH•••N interactions, and **7** has only one. Although this comparison can give some indication of the ability of CH•••N interactions to contribute to the strength of an extended hydrogen bond array, there are many approximations implicit in the study. For example, the solvation of the different hydrogen bond donors and acceptors will vary and also the electrostatic interactions that enhance the hydrogen bond accepting ability of neighbouring pyridine groups should be more pronounced in **5** than in **7** or **8**.

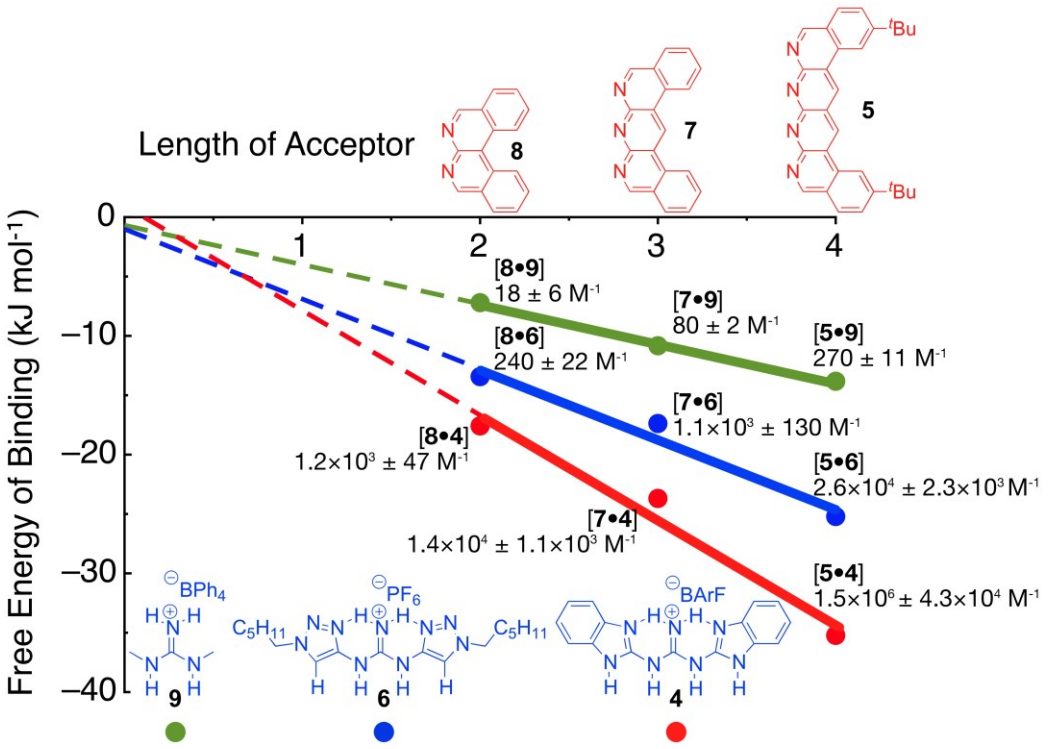


Figure 3.11 | Plot of free energy of acceptor-donor binding vs. the number of hydrogen bonding sites on the acceptor with the appropriate acceptor compounds shown as axis labels. Association constants shown next to points. Repetitions of the binding experiments for each new complex gave K_a's within 12% of the values shown (except 8•9 which was 33%) and the error in data-fitting for each run was <1%. Conditions: 10^{-2} - 10^{-4} M, CH_3CN, 298 K, see section 3.6.1 for details. Used with permission of Dr. Craig C. Robertson.

Despite the limitations of the study, the results show some interesting trends (Figure 3.11). The simple guanidinium derivative **9**, which can only form two NH•••N hydrogen bonds with any of the partners, forms relatively weak complexes in CH_3CN with the association constant only increasing by an order of magnitude across the series **8**→**7**→**5** (K_a = 18-270 M^{-1}). In contrast, bis-triazole **6**, which can form up to two CH•••N interactions in addition to the two NH•••N hydrogen bonds, binds AAAA partner **5** ($K_a = 2.6 \times 10^4$ M^{-1}) two orders of magnitude greater than it binds the 1,8-naphthyridine derivative **8** (K_a = 240 M^{-1}). The peripheral triazole CH groups clearly play a significant role in offering additional stability to the complex through additional hydrogen bonding, although the additional stability from CH hydrogen bonds is lower than from NH hydrogen bond donors (the binding constant for four NH donor **4** increases three orders of magnitude across the series **8**→**7**→**5** in CH_3CN ($K_a = 1.2 \times 10^3$ - 1.5×10^6 M^{-1}). The significant difference in binding strength between **8•9** (K_a = 18 M^{-1}), which features only the central two NH•••N hydrogen

bonds, and **8•6** (K_a = 240 M^{-1}), which has two NH•••N hydrogen bonds plus two secondary CH•••N interactions in an AAAA-DD array, suggests that CH•••N interactions involving triazole groups benefit from secondary electrostatic interactions from adjacent hydrogen bonds in a similar manner to conventional hydrogen bonds involving NH donors. Experimental details and raw data for all acquisitions can be found in Section 3.7.2.

3.5.3 Anion-cation association in H-bond donors

Since the physical properties of guanidines **4**, **6**, and **9** depend on the choice of counterion, some optimisation was necessary – for example, when **6** was paired with $BArF^-$, the resulting salt was an ionic liquid, making clean isolation by precipitation impossible. Since column chromatography always gave rise to partial ion exchange, various weakly-interacting counterions were used in order to isolate the respective guanidinium salts as solids which were still soluble in organic solvents – PF_6^- for **6**, BPh_4^- for **9**, and $BArF^-$ for **4**. Although these are considered weakly-interacting and exhibit non-specific ion pairing only, NMR dilution experiments were carried out in $CDCl_3$ to confirm that changing the counterions would not invalidate any trends drawn with these three compounds. We found that the K_a of the smaller PF_6^- anion with **6** was 480 M^{-1} (Figure 3.12), whereas the large $BArF^-$ anion with **4** was 380 M^{-1} (Figure 3.13), showing that both anions are comparably weak. Since all subsequent UV/vis titrations were carried out at < 1 mM, there would be a negligible fraction bound in each case[†] so the different binding constants calculated here do not need to be factored into subsequent equilibrium models. Similar experiments in CD_3CN failed to observe any binding at all, putting $K_a << 100$ M^{-1} in this solvent.

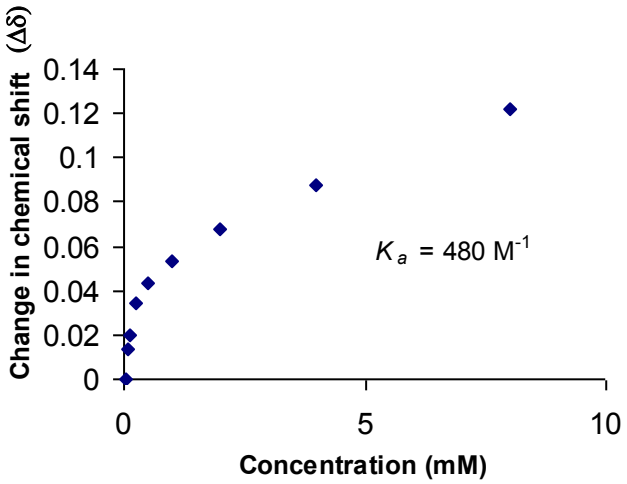


Figure 3.12 | A NMR dilution isotherm for 6 ($CDCl_3$, 298 K, 400 MHz) with concentration ranging from 0.03 to 8 mM.

[†] From crystal structures of the AAA•DDD[21] and AAAA•DDDD[18], the $BArF^-$ and BPh_4^- counterions exhibited non-specific binding in the solid state and would be unlikely to change the acceptor-donor binding constant even if closely associated.

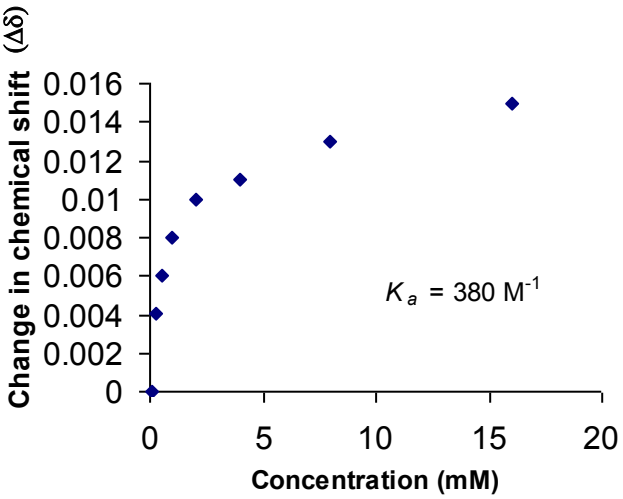


Figure 3.13 | A NMR dilution isotherm for 4 ($CDCl_3$, 298 K, 400 MHz) with concentration ranging from 0.125 to 8 mM.

3.5.4 Acid/base switching of complex formation

As the hydrogen bond donors **4** and **6** utilise a single positive charge to induce preorganisation and display the desired recognition motif, the interaction could in principle be switched off by addition of an appropriate base. Inorganic or anionic bases were ruled out on the basis of poor solubility. Some organic bases have conjugate acids that would be good hydrogen bond donors – they would not need to out-compete **6**, but only the solvent in order to perturb the "resting" spectrum of **5**. Organic bases where the conjugate acid would be a poor hydrogen bond donor are generally less basic than guanidines (such as iPr_2EtN) or basic enough to deprotonate neutral amines, as well (such as *n*-BuLi). 1,8-diazabicyclo5.4.0undec-7-ene (DBU) was chosen as a compromise.

An acid was also needed to reprotonate the neutral **13** back into binding partner **6**, and the conjugate base would need to be a poor hydrogen bond acceptor. Initial attempts with HBF_4 and HPF_6 resulted in complete destruction of the complex, so HI was chosen as iodide is a relatively large, relatively non-specific binder. It was hoped that the DBU hydroiodide generated would largely self-associate and not compete with the hydrogen bonding of **5•6**.

Addition of one equivalent of 1,8-diazabicyclo5.4.0undec-7-ene (DBU) to a solution of **5•6** in $CDCl_3$ cleanly deprotonated the guanidinium group and resulted in the disruption of the strong association of **5** and **6**, as evidenced by shifts of H_b and H_c in the ^1H NMR spectrum to positions consistent with the uncomplexed building blocks (see for example Figure 3.3 for a comparison of complexed and uncomplexed **5**). Reprotonation with one equivalent of HI in CD_3CN smoothly reformed **5•6** (Figure 3.14, experimental details in section 3.7.3). Following this procedure, complex **5•6** was successfully demonstrated to sequentially be switched 'off' and 'on' by successive additions of acid and base (Figure 3.14).

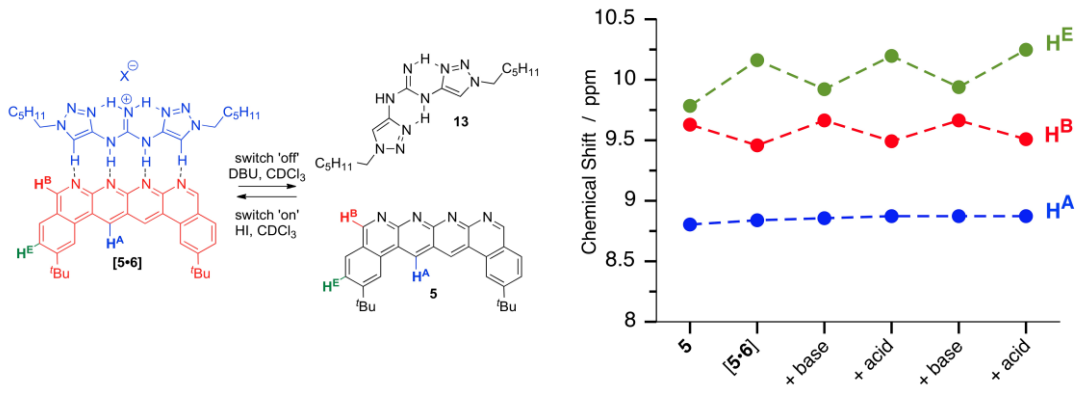


Figure 3.14 | Switching 'off' and 'on' of the AAAA-DDDD$^+$ 5•6 complex by deprotonation by DBU and reprotonation with HI. Counterion X = PF_6^- before the first switching event. The figure shows the change in ^{1}H NMR chemical shift of various protons of 5 before and after several switching events (1 mM, $CDCl_3$, 298 K, 400 MHz). Used with permission of Dr. Craig C. Robertson.

A possibility remained that the spectral changes of **5** during the titration could be due to protonation of one or more of the pyridine nitrogens, and so a sample of **5** was treated with acid (HI, 1 equivalent). The spectral change is shown in Figure 3.14, and is not consistent with the "complexation" or "decomplexation" seen in Figure 3.3.

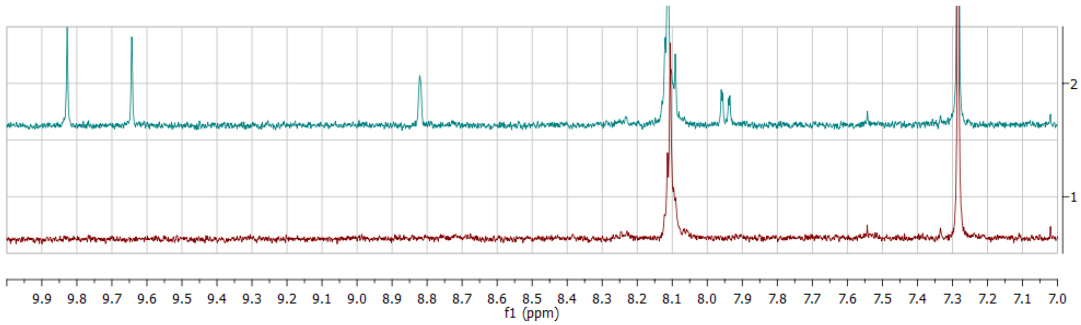


Figure 3.15 | Partial NMR spectra of the aromatic region of 5 (1 mM, 298 K, 400 MHz). Top: neutral species 5. Bottom: 5 + 1eq. HI

3.6 Conclusions

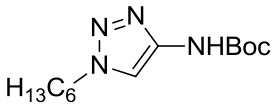
The guanidinium core at the centre of the very strong quadruple H-bonding array **4•5** is responsible for the preorganisation of the DDDD array by internal H-bonding with the adjacent heterocycles. Its design was exploited in the construction of a new quadruple donor array **6**, which utilises two aromatic CH•••N H-bonds in the formation of a new complex **5•6**. The association between the two components was demonstrated to respond to a change in environment when shown to be repeatedly switched 'off' or 'on' by addition of base or acid respectively to the system. The dynamic nature of H-bonding offers the possibility of its implementation into materials with novel functional materials and dynamic properties. Information gleaned from this study is being applied to the incorporation of such multipoint hydrogen bonding arrays into switchable supramolecular polymers.

3.7 Experimental procedures

3.7.1 Synthesis of new compounds

^{1}H and ^{13}C NMR was carried out in $CDCl_3$ at 500/125 or 400/100 MHz. Dry column vacuum chromatography was carried out on silica gel 60 (15-40 μm, Merck) according to the procedure of Harwood[41] and Shusterman.[42]. Low resolution mass spectra were obtained with an Agilent Technologies 1200 LC system with 6130 single quadrupole MS detector, with positive or negative electrospray ionisation. High resolution mass spectra were obtained with a Bruker 3.0 T Apex II Spectrometer, with positive or negative electrospray ionisation. All melting points were determined using a Sanyo Gallenkamp apparatus and are uncorrected. Unless otherwise specified, all reagents were purchased from Acros, Fisher, or Sigma-Aldrich and were used without further purification. The following compounds were prepared according to the given references or used from existing in-house stocks prepared according to the given references: **5** AAAA[18], **7** AAA[20,21], **8** AA [20,21], **4** DDDD[18], bis(2-pyridyl) thiourea.[47] Some compounds had incomplete data available and were included for general interest but would require re-analysis if remade: **14, 15, 17-19.**

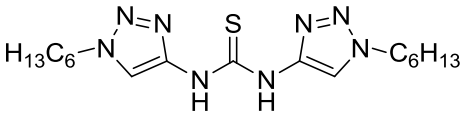
Tert-butyl _N_-(1-hexyl-1_H_-1,2,3-triazol-4-yl)carbamate (11)



To a solution of 1-hexyl-1*H*-1,2,3-triazole-4-carboxylic acid **10**[6] (5.91 g, 30 mmol) in *t*-butanol (50 ml) and triethylamine (4.46 ml, 1.07 equivalents) was added diphenylphosphoryl azide (7.23 ml, 1.03 equivalents). The resulting solution was heated to 100 °C for 16 h. The solution was cooled and the solvent removed under reduced pressure. The residue was dissolved in ethyl acetate (300 ml) and washed with saturated aqueous sodium carbonate solution (2 × 200 ml), hydrochloric acid solution (0.1 M, 2 × 200 ml), and brine (200 ml). The organic layer was treated with magnesium sulfate and activated carbon before filtering through a short plug of silica

gel (standard "wet flash" silica, 40-63 μm particle size). The solvent was removed under reduced pressure to give a dark oil (6.7 g) which was purified by dry column vacuum chromatography (20% ethyl acetate in heptane) to give the title compound as a white solid (2.77 g) in 33% yield. mp 124 - 129 °C (ethyl acetate/heptane); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (s, 1H), 7.33 (s, 1H), 4.30 (m, 2H), 1.53 (m, 2H), 1.32 (m, 9H), 0.91 (m, 6H), 0.90 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 152.2, 144.2, 111.4, 81.2, 50.8, 31.2, 30.0, 28.3, 26.1, 22.4, 13.9. MS: 213 ($M-C_4H_9 + H^+$, 80), 269 ($M+H^+$, 100), 291 ($M+Na^+$, 10), 559 ($2M + Na^+$, 30). HRMS $M+Na^+$ 291.1783. $C_{13}H_{24}N_4O$ requires 291.1792.

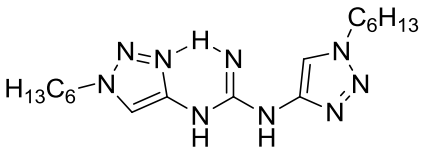
1,3-Bis(1-hexyl-1H-1,2,3-triazol-4-yl)thiourea (12)



To a solution of **11** (528 mg, 1.97 mmol) in CH_2Cl_2 (20 ml) was added trifluoroacetic acid (5 ml), and the reaction stirred at 25 °C for 16 h. The solvent and excess trifluoroacetic acid were removed under reduced pressure to give a mixture of the free amine and the trifluoroacetamide. The crude mixture was dissolved in a saturated solution of ammonia in methanol (7 M, 10 ml) and stirred for 6 days. The solvents were removed under reduced pressure and the resulting crude material was dissolved in pyridine (10 ml) and carbon disulfide (2.5 ml). The solution was heated to reflux with an oil bath temperature of 130 °C for 16 h. The solution was then cooled and the solvents removed under reduced pressure. The residue was dissolved in ethyl acetate (50 ml) and washed with saturated aqueous sodium carbonate solution (50 ml) and water (2 × 50 ml). The organic phase was dried over magnesium sulfate and concentrated under reduced pressure to give an orange solid (1.22 g) which was purified by dry column vacuum chromatography (25% ethyl acetate in heptane) to give the title compound as an orange solid (274 mg) in 83% yield over two steps. mp 119-123 °C (ethyl acetate/heptane); ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.5-9.5 (br.s, 2H), 9.0-8.0 (br.s, 2H), 4.38 (m, 4H), 1.95 (m, 4H), 1.34 (m, 12H), 0.90 (m, 6H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 175.0, 145.3, 114-112 (br) 51.1, 31.1, 30.1, 26.1, 22.4,

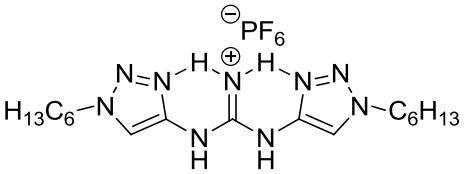
14.0. MS-ESI$^+$: 379 (M+H$^+$, 100), 779 (2M + Na$^+$, 30); MS-ESI$^-$: 377 (M-H$^-$, 100). HRMS M+H$^+$ 379.2373. $C_{17}H_{31}N_8S$ requires 379.2387.

1,3-Bis(1-hexyl-1H-1,2,3-triazol-4-yl)guanidine (13)



To a solution of **12** (172 mg, 0.456 mmol) in chloroform (2.5 ml) was added mercury oxide (138 mg, 1.4 equivalents) and ammonia in methanol (7 M, 2.5 ml). The solution was stirred at 20 °C for 2 h, then filtered through a plug of celite and the solvent and excess ammonia removed under reduced pressure to give the title compound as a white solid (150 mg) in 91% yield. mp 139-142 °C (chloroform/methanol); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.0-7.5 (br.s, 2H), 7.32 (s, 2H), 4.30 (m, 4H), 1.90 (m, 4H), 1.33 (m, 12H), 0.90 (m, 6H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 151.4, 151-150 (br), 112.2, 50.7, 31.2, 30.1, 26.1, 22.4, 14.0. MS: 362 (M+H$^+$, 100), 723 (2M+H$^+$, 5). HRMS M+H, ^{13}C isotope peak$^+$ 363.2789. $C_{16}{}^{13}C_1H_{31}N_9$ requires 363.2809.

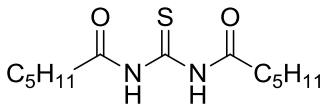
1,3-Bis(1-hexyl-1H-1,2,3-triazol-4-yl)guanidinium hexafluorophosphate (6)



To a suspension of **6** (40 mg, 0.11 mmol) in aqueous acetic acid (8 M, 6 ml) in a glass screwcap vial was added a solution of sodium hexafluorophosphate (37 mg, 2.0 equivalents). The vial was sealed and violently agitated for 2 h, with sonication for 5 minute intervals every 30 minutes. The product was isolated as a single white pellet, which was crushed on a filter and washed with copious quantities of water, then dried for 16 h in a vacuum oven (40 °C, 1 millibar) to give the title compound as a white solid (38 mg) in 70% yield. mp 143-145 °C (aqueous acetic acid); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.61 (s, 2H), 4.36 (m, 4H), 1.92 (m, 4H), 1.33 (m, 12H), 0.90 (m,

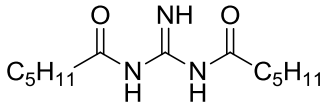
6H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 180.8, 144.0, 112.5, 51.3, 31.1, 30.0, 26.0, 22.4, 13.9. ^{19}F NMR (376 MHz, $CDCl_3$) 70.55 (d, ^1J = 714 Hz). MS-ESI$^+$: 362 ($M+H^+$, 100); MS-ESI$^-$: 145 (PF_6^-, 100), 360 ($M-H^-$, 15). HRMS $M+H^+$ 362.2768. $C_{17}H_{32}N_9$ requires 362.2775.

N,*N*′-Carbonothioylbishexanamide (17)



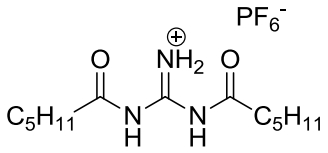
Hexanoyl chloride (0.29 mL, 283 mg, 2.1 mmol) was added dropwise to a solution of thiourea (78 mg, 1 mmol) and triethylamine (0.292 mL, 212 mg, 2.1 mmol) in tetrahydrofuran (5 mL) under a nitrogen atmosphere at 5 °C. The reaction mixture was stirred at room temperature for 1 hour, then poured into saturated aqueous sodium carbonate (20 mL) and extracted with ethyl acetate (2 × 20 mL). The combined organic fractions were dried over magnesium sulfate and dried *in vacuo*, then purified by dry column vacuum chromatography (10% ethyl acetate, 90% heptane) to give **17** as a yellow oil (175 mg, 64%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 2.48 (m, 2H), 1.72 (m, 2H), 1.36 (m, 4H), 0.93 (m, 3H). ^{13}C NMR (125 MHz, $CDCl_3$): δ = 177.1, 172.8, 38.2, 31.1, 24.2, 22.3, 13.9.

N,*N*′- Carbonimidoylbishexanamide (18)



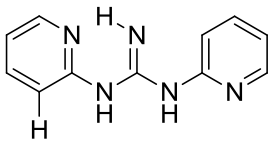
Compound **17** (174 mg, 0.64 mmol) was dissolved in chloroform (3 mL) and methanolic ammonia (3 mL, 7 M) and stirred with mercury oxide (195 mg, 1.4 equivalents) for 16 h at room temperature. The mixture was filtered through celite, and evaporated to give **18** (135.8 mg, 82%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 2.24 (m, 2H), 1.66 (m, 2H), 1.35 (m, 4H), 0.92 (m, 3H).

N,*N*′- Carbonimidoylbishexanamidium hexafluorophosphate (14)



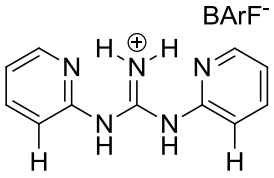
Material produced using a similar protocol to compound **6**. ^{1}H NMR (400 MHz, $CDCl_3$): δ = 2.63 (m, 2H), 1.75 (m, 2H), 1.36 (m, 4H), 0.93 (m, 3H).

1,3-Dipyridin-2-ylguanidine (19)



N,*N'*-Bis(2-pyridyl)thiourea (500 mg, 2.13 mmol)[47] was dissolved in chloroform (10 mL) and methanolic ammonia (10 mL, 7 M) and stirred with mercury oxide (650 mg, 1.4 equivalents) for four days at room temperature. The mixture was filtered through celite, and evaporated to give **19** as a colourless solid (368.5 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.23 (m, 2H), 7.55 (m, 2H), 6.83 (m, 4H).

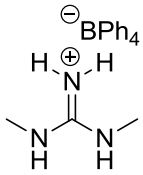
1,3-Dipyridin-2-ylguanidinium tetrakis(3,5-trifluoro-methyl)phenylborate (15)



Compound **19** (278 mg, 1.28 mmol) was dissolved in aqueous acetic acid (8 M, 20 mL) and the resultant solution filtered. To this was added a filtered solution of sodium tetrakis(3,5-trifluoromethyl)phenylborate (1.13 g, 1.28 mmol) dissolved in a minimum amount of acetic acid (8 M, aq.) and the solution stirred at room temperature until no further precipitate formed. The precipitate was collected on celite and washed with copious water, then CH_2Cl_2. The organic washings were

concentrated under reduced pressure and purified by dry column vacuum chromatography (5% methanol and 0.5% triethylamine in chloroform) to give compound **15** as a colourless solid (220 mg, 44%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.27 (m, 2H), 7.70 (m, 8H), 7.50 (m, 2H), 7.17 (m, 2H), 7.09 (m, 4H).

1,1-Dimethylguanidine tetraphenylborate (9)



To a solution of 1,1-dimethylguanidine sulfate (2.00 g, 7.35 mmol) in aqueous acetic acid (8 M, 20 ml) was added a solution of sodium tetraphenylborate (3.00 g, 8.77 mmol) in aqueous acetic acid (8 M, 20 ml) dropwise with stirring. The reaction was stirred at 25 °C for 1 hour, then filtered, washed with copious water, and dried for 16 h in a vacuum oven (40 °C, 1 millibar) to give the title compound as a white, highly electrostatic solid (3.05 g) in 98% yield. mp 204-208 °C (aqueous acetic acid); ^{1}H NMR (400 MHz, CD_3CN) δ 7.32 (m, 8H), 7.04 (m, 4H), 6.88 (m, 4H), 5.83 (br.s, 4H), 2.95 (s, 6H). ^{13}C NMR (101 MHz, CD_3CN) δ 164.7 (m), 136.6, 126.5 (m), 122.7, 38.7. MS: 87 (M^+, 100), 88 ($M+H^+$, 10). HRMS M^+ 88.0870. $C_3H_{10}N_3^+$ requires 88.0870.

3.7.2 Evaluating binding constants in MeCN

Experimental procedures were the same as those described in Chapter II (Section 2.7.2). Some of the data from Chapter II is reproduced here for consistency, since all the data required to create Figure 3.11 is reproduced here in Table 3.1. For simplicity, only the binding pair is reported; in each case the 1:1 equilibrium constant is reported (as an energetic value in kJ mol^{-1})

Binding Pair	Concentration	Titration 1 K_a (kJ mol^{-1})	Titration 2 K_a (kJ mol^{-1})	Titration 3 K_a (kJ mol^{-1})	Average K_a (kJ mol^{-1})	Standard deviation
5•9[(d)]	221 μM	-13.92	-13.71	-13.85	-13.82	0.107
7•9[(d)]	562 μM	-10.83	-10.75	-10.90	-10.83	0.073
8•9[(d)]	422 μM	-8.05	-6.52	-7.02	-7.20	0.779[(b)]
5•6[(d)]	212 μM	-25.42	-25.22	-24.98	-25.21	0.222
7•6[(c)(d)]	1210 μM	-17.53	-17.56	-17.03	-17.38	0.237
8•6[(d)]	1750 μM	-13.16	-13.56	-13.58	-13.43	0.236
5•4[(a,c)]	3.6 μM	-35.40	-35.81	-34.45	-35.22	0.698
7•4[(a)]	104 μM	-23.66	-23.51	-23.89	-23.69	0.192
8•4[(d)]	2050 μM	-17.58	-17.57	-17.74	-17.63	0.094

Table 3.1 \| Three repetitions of binding energy determination from three separate titrations, the average value used in discussion, and the standard deviation for each set. Each value was recorded in MeCN at 298 K and the given concentration with a 10.0 mm path length unless otherwise specified. (a) from chapter II (b) Solubility of the components restricted the maximum concentration and partial overlapping spectra meant that the error was unavoidably high. (c) a 2:1 binding mode was also visible and was separated from the 1:1 binding mode by the fitting software. (d) a cell with a 2.00 mm path length was used.

The stoichiometry of each experiment was determined by a modified Job method[44] of replotting the existing data, which cut down on experimental complexity. The validity of the fitting model was also verified by the presence of small, randomised residuals in the fitting plot; a systematic bias in the residuals that peaked at 0.5 eq. donor added was observed in the cases where a 2:1 acceptor:donor equilibrium was present but modelled as a 1:1. Reproduced below is an example of the data fitting process for **5•6**, showing the ReactLab I/O spreadsheet (Figure 3.17), the ReactLab working window

(Figure 3.18), and UV/vis (Figure 3.19) and NMR (Figure 3.20) Job plots showing a 2:1 stoichiometry. Similar analysis was carried out for each data point given in Table 3.1.

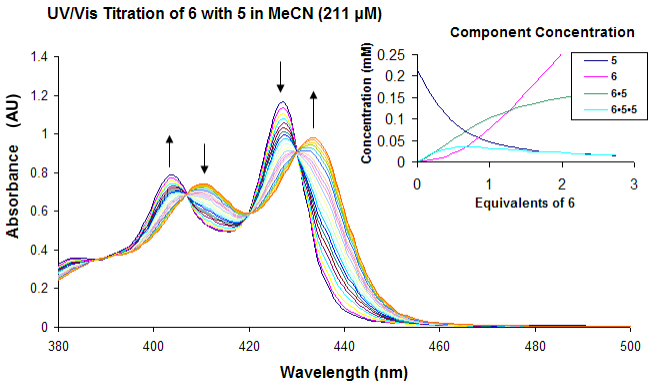


Figure 3.16 | UV/vis titration of 6 with 5 in MeCN. Inset: a profile of component stoichiometry determined from full-spectral multivariate analysis in he ReactLab Equilibrium package.

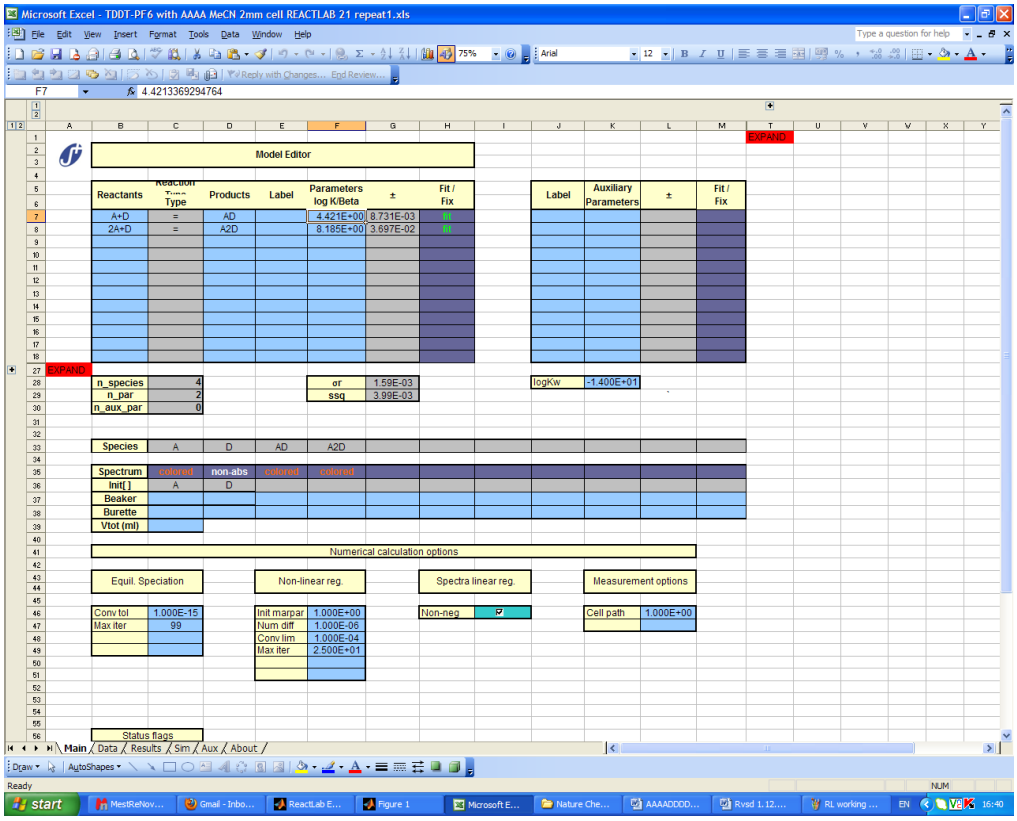


Figure 3.17 | UV/vis titration of 6 with 5 in MeCN, ReactLab Input/Output displaying log(K) = 4.421 for the 6•5 equilibrium and log(K) = 8.185 for the 5·(6·5) equilibrium. ReactLab can output values of sequential binding as Kn or βn values depending on the model used; here we obtain the value in the format β_n.

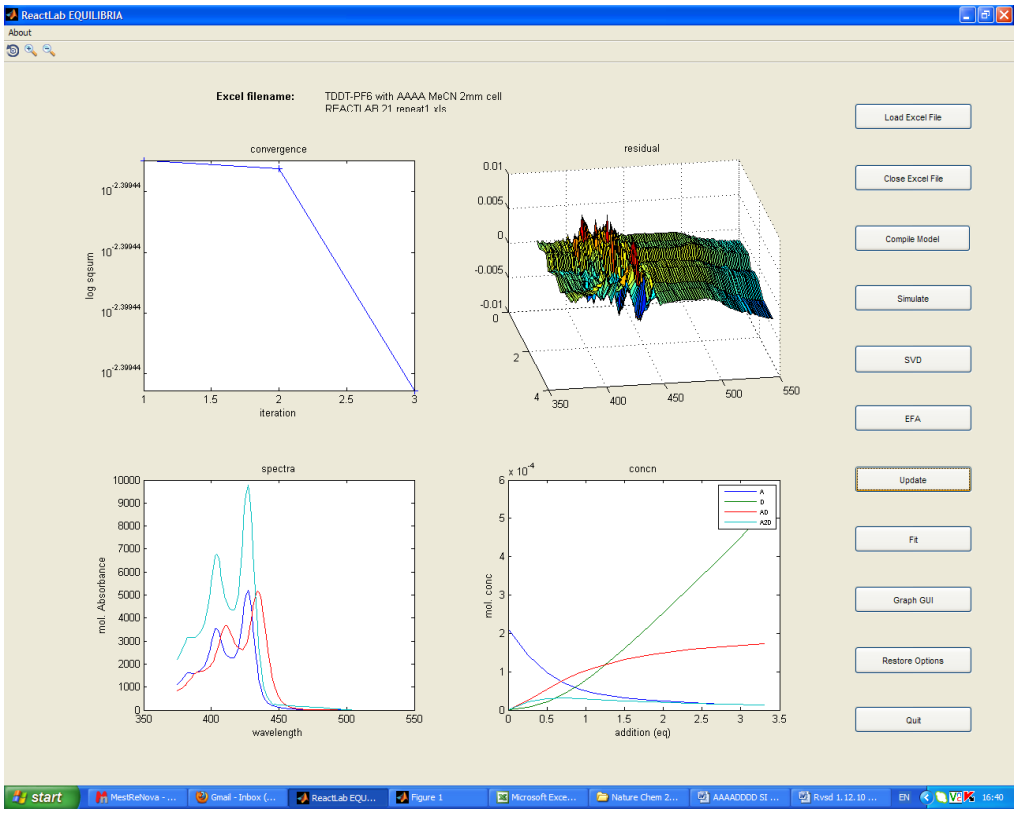


Figure 3.18 | UV/vis titration of 6 with 5 in MeCN, ReactLab Working Window. Top left, plot of least-squares optimisation during a fit. Top right, a 3D plot of the residuals for the fit. Bottom left, the calculated spectrum for each species. Bottom right, concentration profile for the species in solution.

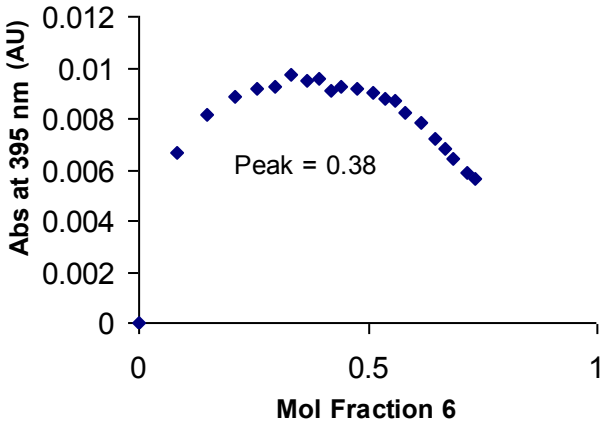


Figure 3.19 | A Job Plot for the titration of 6 with 5, showing a mix of 1:1 and 2:1 equilibria (UV/vis, 0.2 mM, 298 K, 395 nm)

At most UV/vis titration wavelengths, only a 1:1 stoichiometry of **6** to **5** was apparent. Only a few selected wavelengths such as 395 nm (Figure 3.19) showed a deviation

from a 1:1 equilibrium (e.g. a peak at mol fraction 0.5) , so as confirmation an NMR titration was carried out, adding aliquots of **6** into a solution of **5**, and an isotherm was obtained that clearly shows a 2:1 equilibrium (**6•5•5**, Figure 3.20).

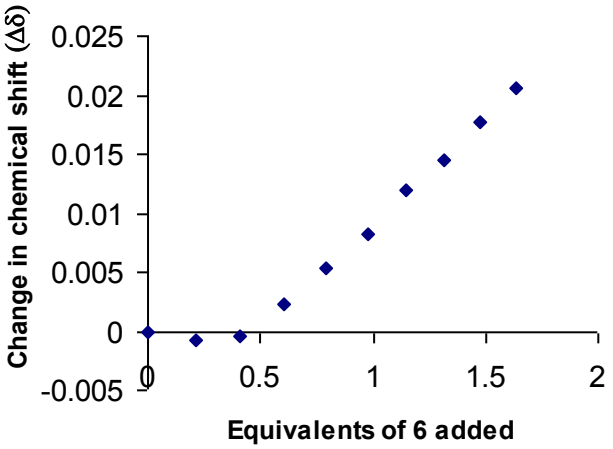


Figure 3.20 | An NMR titration isotherm for 6 being titrated into a fixed concentration of 5, showing a sharp change at 0.5 equivalents added, consistent with a 6•5•5 equilibrium (1 mM, 298 K, 400 MHz)

3.7.3 Evaluating acid/base switching

Complex **5•6** could be switched on and off by deprotonating and protonating the guanidinium motif within **5**. This could be tracked by ^{1}H NMR spectroscopy.

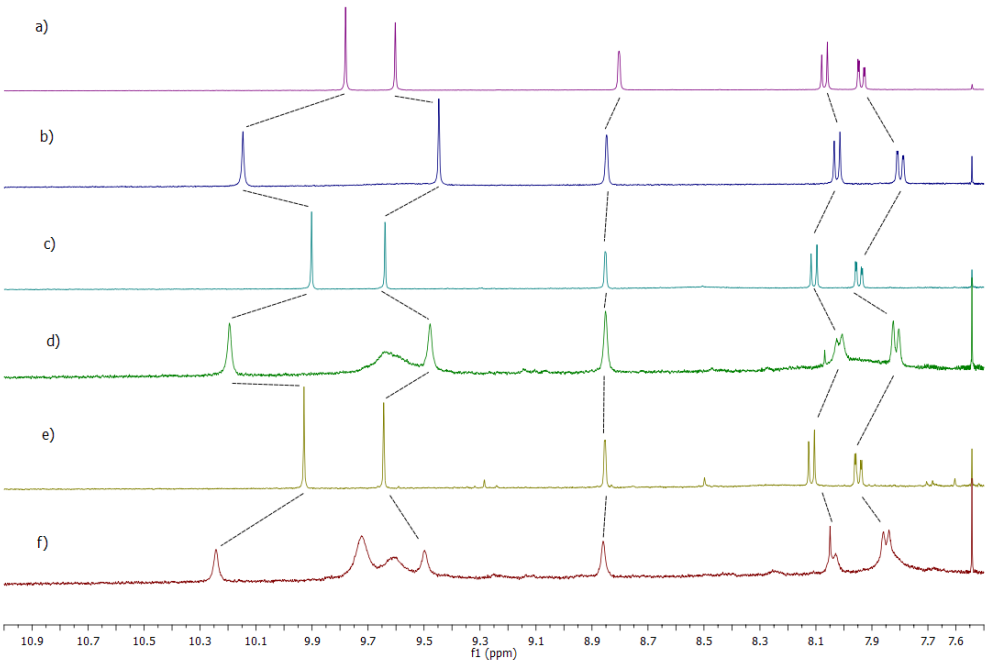


Figure 3.21 | The aromatic region of the NMR spectrum corresponding to the switching of complex 5•6 with successive aliquots of acid and base. (1 mM, 298 K, 400 MHz)

Figure 3.21 shows the full NMR spectrum corresponding to Figure 3.14. The experiment was carried out as follows: stock solutions of HI (100 mM in CD_3CN) and DBU (100 mM in $CDCl_3$) were prepared from 57% aqueous HI, and neat DBU.

A 1 mM solution of **5** in 1.0 mL of $CDCl_3$ was analysed by NMR spectroscopy (Spectrum b), then 100 µL of a 10 mM solution of **6** in $CDCl_3$ was added and a small amount of $CDCl_3$ evaporated with dry nitrogen so that the NMR tube contained 1.0 mL of a 1 mM solution of **5•6** (Spectrum b). To this was added 10 µL of the DBU stock solution (spectrum c), followed by 10 µL of the HI stock solution (spectrum d). The base and acid cycle was repeated once more (spectrums e and f). The increasing broadness of the successive spectra is attributed to DBU•HI being soluble in $CDCl_3$ and weakly binding to **5** *via* monodentate hydrogen bonding.

3.8 References

(1) Beijer, F. H.; Kooijman, H.; Spek, A. L.; Sijbesma, R. P.; Meijer, E. W. *Angew. Chem. Int. Ed.* **1998**, *37*, 75–78.

(2) Beijer, F. H.; Sijbesma, R. P.; Kooijman, H.; Spek, A. L.; Meijer, E. W. *J. Am. Chem. Soc.* **1998**, *120*, 6761–6769.

(3) Sijbesma, R. P.; Meijer, E. W. *Chem. Commun.* **2003**, 5–16.

(4) Kuykendall, D. W.; Anderson, C. A.; Zimmerman, S. C. *Org. Lett.* **2009**, *11*, 61–64.

(5) Hisamatsu, Y.; Shirai, N.; Ikeda, S.-i.; Odashima, K. *Org. Lett.* **2009**, *11*, 4342–4345.

(6) Hisamatsu, Y.; Shirai, N.; Ikeda, S.; Odashima, K. *Org. Lett.* **2010**, *12*, 1776–1779.

(7) Corbin, P. S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9710–9711.

(8) Zimmerman, S. C.; Murray, T. J. *Phil. Trans. R. Soc. A* **1993**, *345*, 49–56.

(9) Bosman, A. W.; Sijbesma, R. P.; Meijer, E. M. *Mater. Today* **2004**, *7*, 34–39.

(10) Ligthart, G. B. W. L.; Ohkawa, H.; Sijbesma, R. P.; Meijer, E. W. *J. Am. Chem. Soc.* **2005**, *127*, 810–811.

(11) Fox, J. D.; Rowan, S. J. *Macromolecules* **2009**, *42*, 6823–6835.

(12) Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. *Science* **1997**, *278*, 1601–1604.

(13) Lehn, J.-M. *Polym. Int.* **2002**, *51*, 825–839.

(14) Aida, T.; Meijer, E. W.; Stupp, S. I. *Science* **2012**, *335*, 813–817.

(15) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.

(16) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810–2819.

(17) Park, T.; Zimmerman, S. C.; Nakashima, S. *J. Am. Chem. Soc.* **2005**, *127*, 6520–6521.

(18) Blight, B. A.; Hunter, C. A.; Leigh, D. A.; McNab, H.; Thomson, P. I. T. *Nature Chem.* **2011**, *3*, 244–248. (b) Wilson, A. J. *Nature Chem*. **2011**, 3, 193–194.

(19) Murray, T. J.; Zimmerman, S. C.; Kolotuchin, S. V. *Tetrahedron* **1995**, *51*, 635–648.

(20) Djurdjevic, S.; Leigh, D. A.; McNab, H.; Parsons, S.; Teobaldi, G.; Zerbetto, F. *J. Am. Chem. Soc.* **2007**, *129*, 476–477.

(21) Blight, B. A.; Camara-Campos, A.; Djurdjevic, S.; Kaller, M.; Leigh, D. A.; McMillan, F. M.; McNab, H.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2009**, *131*, 14116–14122.

(22) Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 1039–1057.

(23) Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1009–1020.

(24) Lehn, J.-M. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 89–112.

(25) Pedersen, C. J. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1021–1027.

(26) Desiraju, G. R. *Angew. Chem. Int. Ed.* **2011**, *50*, 52–59.

(27) Desiraju, G. R. *Acc. Chem. Res.* **2002**, *35*, 565–573.

(28) Desiraju, G. R. *Acc. Chem. Res.* **1991**, *24*, 290–296.

(29) Allerhand, A.; Von Rague Schleyer, P. *J. Am. Chem. Soc.* **1963**, *85*, 1715–1723.

(30) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063–5070.

(31) Sutor, D. J. *Nature* **1962**, *195*, 68–69.

(32) Desiraju, G. R. *Chem. Commun.* **2005**, 2995–3001.

(33) Biradha, K. *CrystEngComm* **2003**, *5*, 374–384.

(34) Desiraju, G. R.; Panigrahi, S. K. *Prot. Struct. Funct. Bioinf.* **2007**, *67*, 128–141.

(35) Stahl, M.; Bissantz, C.; Kuhn, B. *J. Med. Chem.* **2010**, *53*, 5061–5084.

(36) Bianchi, A.; Bowman-James, K.; García-España, E. *Supramolecular Chemistry of Anions*; Wiley-VCH: New York, 1997.

(37) Sessler, J. L.; Gale, P. A.; Cho, W.-S. *Anion Receptor Chemistry*; Royal Society of Chemistry: Cambridge, UK, 2006.

(38) McDonald, K.; Hua, Y.; Flood, A.; Gale, P. A.; Dehaen, W. *Top. Het. Chem.* **2010**, 24, 341–366.

(39) Hua, Y.; Flood, A. H. *Chem. Soc. Rev.* **2010**, *39*, 1262–1271.

(40) Cook, J. L.; Hunter, C. A.; Low, C. M. R.; Perez-Velasco, A.; Vinter, J. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 6275–6277.

(41) Harwood, L. M. *Aldrichim. Acta* **1985**, *18*, 25.

(42) Shusterman, A. J.; McDougal, P. G.; Glasfeld, A. *J. Chem. Ed.* **1997**, *74*, 1222.

(43) JPlus Consulting, http://www.jplusconsulting.com accessed on 10th January 2012.

(44) MacCarthy, P. Anal. Chem. 1978, 50, 2165.

(45) IDCalc, http://proteome.gs.washington.edu/software/IDCalc/ accessed on 6th March 2012.

(46) Bastings, M. M. C.; de Greef, T. F. A.; van Dongen, J. L. V.; Merkx, M.; Meijer, E. W. *Chem. Sci.*. **2010**, *1*, 79–88.

(47) Saxena, A.; Pike, R. D.; *J. Chem. Crystallogr.*. **2007**, *3*, 755-764.

Chapter IV

Towards a Supramolecular Polymer Containing an Extremely Strong, pH-Responsive AAA-DDD Motif

Acknowledgements

The following people are gratefully acknowledged for their contribution to this chapter: Mr Malcolm Gall synthesised and characterised compounds **61-65**. Dr Barney Walker developed the synthetic protocol for compounds **66**, **59**, and **39** and synthesised compounds **66** and **59**. Dr Barney Walker also carried out the experimental work in Scheme 4.5 and Scheme 4.12. Abidin Balan and Marko Nieuwenhuizen (Sijbesma group) carried out the concentration studies reported in Figure 4.13 and section 4.5.2. Dr Barney Walker and Prof. Tim Swager (MIT) carried out the solar cell and organic semiconductor studies in Section 4.5.3.

4.1 Synopsis

This chapter concerns the synthesis of triply hydrogen bonded motifs containing only positive secondary interactions, that is, all of the hydrogen bond acceptors (A) are on one molecule and all the donors (D) are on the other. The AAA and DDD motifs are based on a design previously developed in our group but functionalised with an azide group. The azide is tethered by a copper-catalysed alkyne-azide cycloaddition, allowing incorporation into supramolecular polymers. Unsuccessful attempts were made to incorporate the supramolecular polymer into organic electronic materials. Furthermore, the DDD motif is pH-responsive and can be switched on with acid in a similar way to the DDDD motif presented in Chapter III. The binding constant of the tethered AAA and DDD motifs is the same as the untethered systems, at a very strong $3.8 \times 10^{10}\ M^{-1}$ *in* $CDCl_3$.

4.2 Introduction to responsive supramolecular polymers

A supramolecular polymer is defined as any polymer based on monomers held together by directional and reversible secondary (non-covalent) interactions. Meijer's Q&A in Nature in 2008 serves as an excellent primer,[1] and several comprehensive reviews also exist on this topic.[2-4] Because the polymer backbone is composed of supramolecular interactions, the polymer properties depend on the same factors that affect hydrogen bonding. In this respect, all supramolecular polymers are stimuli-responsive in the sense that a change of solvent, concentration or temperature will alter the degree of polymerisation of a supramolecular polymer (Figure 4.1), but supramolecular polymers responsive to chemical, photonic, or electrochemical stimuli are somewhat less well-known. In this introduction we will present a highlight of some previous work which encompasses an introduction to supramolecular polymers, and the switching thereof.

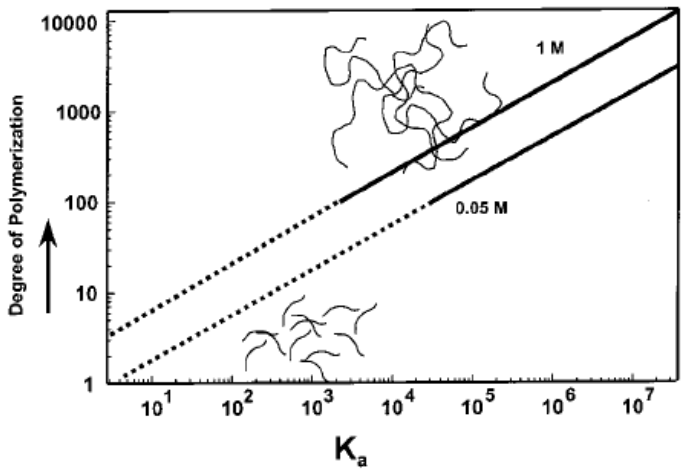


Figure 4.1 | The relationship between polymer length (DP), binding constant and concentration.[2]

The use of hydrogen bonding as a backbone interaction in a supramolecular polymer can be traced to the group of Lehn,[5] who used an ADA-DAD uracil-diamidopyridine motif with a modest binding constant of 90 M^{-1}.[6] Nevertheless, his "rigid rods" exhibited behaviours characteristic of a supramolecular polymer, as opposed to previous systems based on one or two hydrogen bonds. (Figure 4.2).

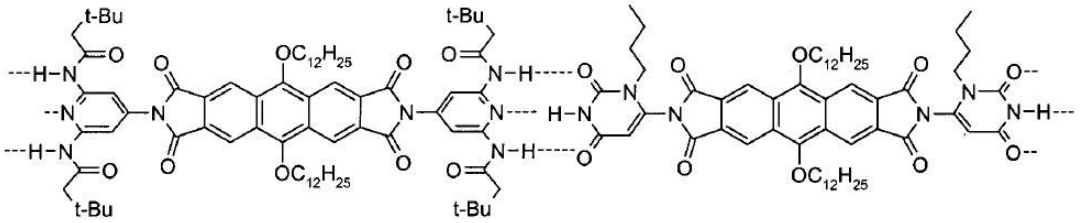


Figure 4.2 | Lehn's rigid rods, as an alternating copolymer held together by an ADA•DAD recognition between an diamidopyridine and uracil motifs.[5]

The biggest development in the field of supramolecular polymers came a few years later, with Meijer and Sijbesma's report[7] of a class of urea-substituted pyrimidinones colloquially referred to as UPy, a self-complementary AADD motif with a K_a of 6×10^7 M^{-1}. The motif can be synthesised in one step from commercial or readily-available materials, and trivially incorporated into ditopic (double-ended) monomers that self-assemble to polymeric systems (Figure 4.3). UPy has since been incorporated into polymers with low melt temperatures,[8] self-healing polymers,[9] molecular cages[10] and a wide variety of other applications (Chapter I, Section 1.3.3).

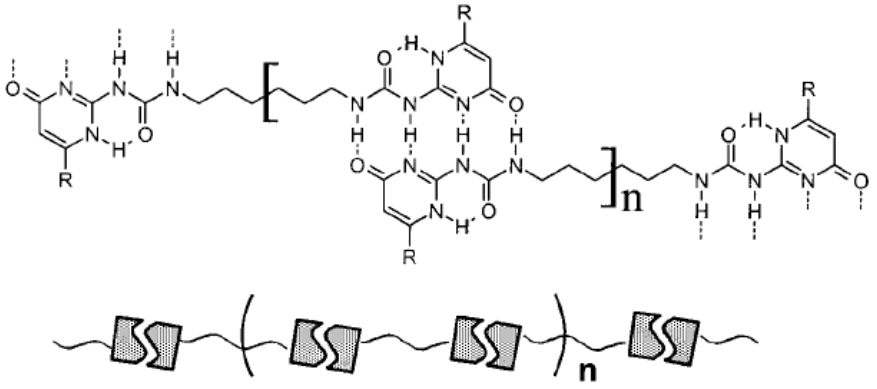


Figure 4.3 | UPy, shown self-associated as a ditopic monomer. R = $C_{12}H_{25}$.[11]

UPy is particularly interesting because the intramolecular interaction between ditopic monomers can be disrupted by the addition of extra monomers, which compete and induce a "chain capping" effect at low doping levels. This is a disadvantage for the

purposes of synthesis and purification, as only a few percent of an unreacted monomer can cause a significant drop in polymer chain length (Figure 4.4).[7]

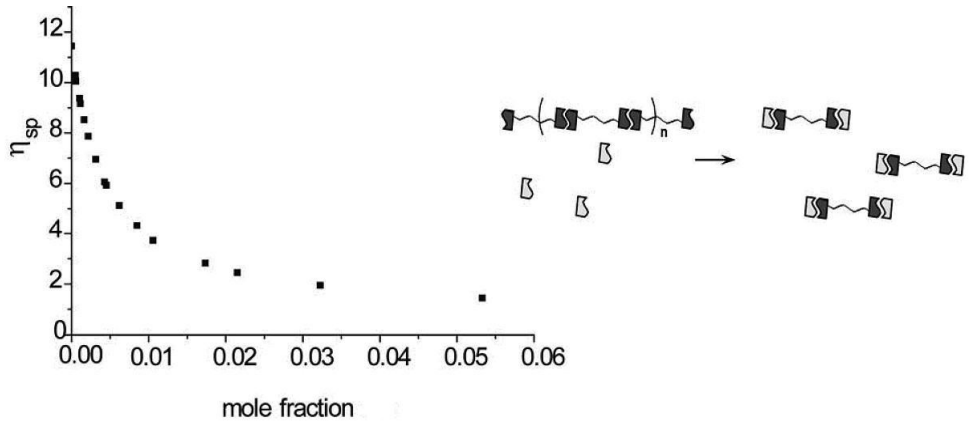


Figure 4.4 | the specific viscosity, and therefore the polymer chain length, of UPy polymers markedly drops when less than 1% of monotopic monomer is added as a capping agent.[2]

However, this effect can be exploited to cause on-demand depolymerisation through photo-induced dissociation of a masked monomer.[12] The masked monomer **1** is dispersed through a solution or polymer blend containing UPy linkages. Due to steric hindrance and the poor hydrogen bond accepting ability of benzyl ethers, **1** does not significantly hydrogen bond to unmodified UPy. Upon irradiation with UV light, the 2-nitrobenzyl alcohol moiety cleaves to furnish 4-hydroxypyrimidine, which immediately tautomerises to **2**, which presents a UPy motif. This creates a large drop in viscosity caused by the disruption of the supramolecular polymer backbone.

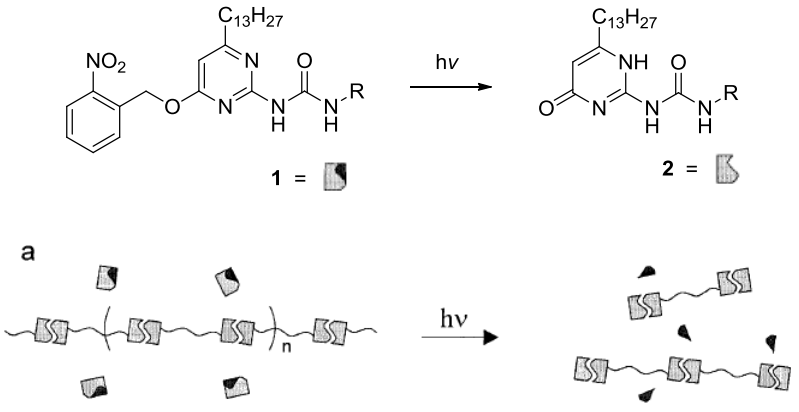


Figure 4.5 | UPy 2 can be released by UV irradiation of masked precursor 1.[12]

The light-initiated depolymerisation is irreversible, making this an unsuitable system for repeated switching applications.

An alternative quadruply hydrogen bonded motif is Zimmerman's diaminonaphthyridine.[13,14] This DAAD motif is slightly weaker than UPy, but is not self-complementary (binding to an ADDA motif). Because of its recognition fidelity, it can be used to selectively crosslink blends of normally-immiscible conventional polymers of 15-30 kDa by decorating different monomers with different recognition motifs and inducing crosslinking on mixing (Figure 4.6) The crosslinking is evident in the marked increase in viscosity of the crosslinked polymers when compared to the un-crosslinked species and can be monitored by the so-called "inverted vial test" wherein a gelator defies gravity (e.g. Figure 4.7c).

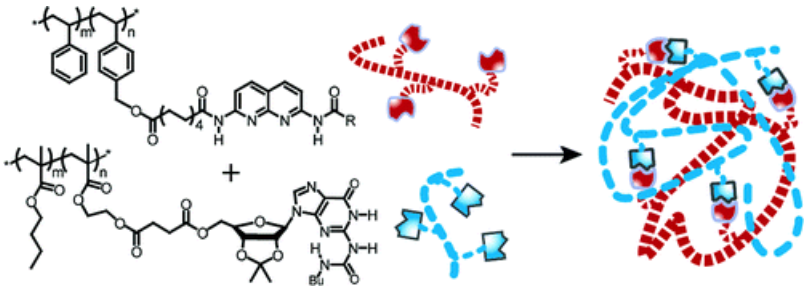


Figure 4.6 | Polystyrene and poly(butyl)methacrylate are normally immiscible, but when decorated with complementary recognition motifs, will crosslink.[14]

A particularly interesting development with stimuli-responsive supramolecular polymers came in 2011 when Zimmerman developed a DAAD motif that can be switched with redox chemistry (Figure 4.7).[15] When reduced, the well-fitting motif **3•4** has strong binding of K_a = 1.1 × 10^6 M^{-1}. When oxidised, however, the binding **5•4** drops to K_a = 6.7 x 10^2 M^{-1}.

The 2,000-fold decrease in binding constant can be used to selectively inhibit a previously-developed cross-linked gel (Figure 4.6), by adding some of the redox-responsive monomer.

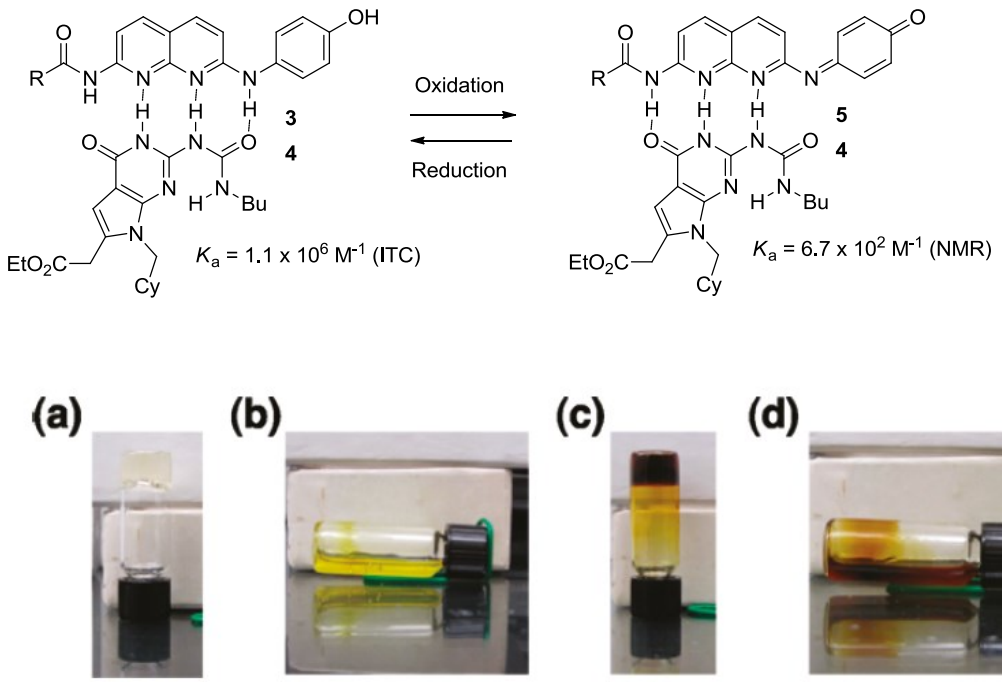


Figure 4.7 | Upon oxidation, the hydrogen bond donating NH of 3 becomes a hydrogen bond accepting imine in 5, lowering the binding to 4 by over 3 orders of magnitude. (a) polymer blend showing crosslinking (b) blend with added 3 (c) reformation of the gel by oxidation of 3 to 5 (d) re-destruction of the gel by reduction of 5 back to 3. [15]

In the default reduced state, **3** is a good competitor for the crosslinking and the gel collapses (Figure 4.7, **a** to **b**). When the solution was oxidised with salcomine/O_2, the gel was reformed because the oxidised **5** is now a poor match for the intra-strand recognition (Figure 4.7, **c**). Finally, re-reduction of **5** to **3** was effected by hydroquinone, leading to collapse of the gel again (Figure 4.7, **d**). This system could in principle be extended to a supramolecular polymer capable of being cycled between gel and liquid multiple times, under the influence of oxidation and reduction.

4.3 Design and synthesis of tethered recognition motifs

4.3.1 Towards tethered recognition motifs

Incorporation of the recognition motif of our existing work (Figure 4.8) into a supramolecular polymer has been a valuable target in the group for some time, but efforts until now were unsuccessful.

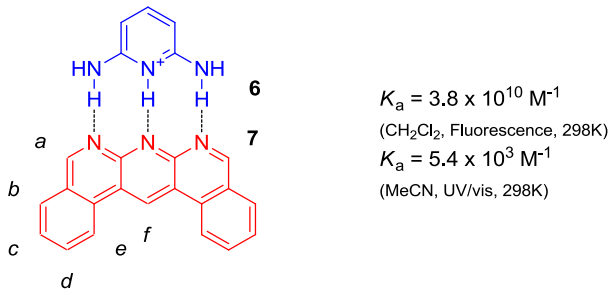
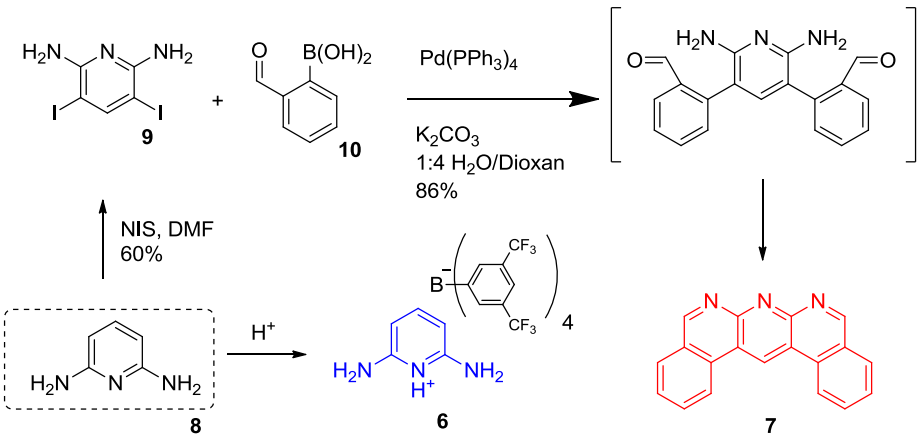


Figure 4.8 | The previously-developed[16,17] triple hydrogen bonded array AAA•DDD, binding in CH_2Cl_2 and CH_3CN.

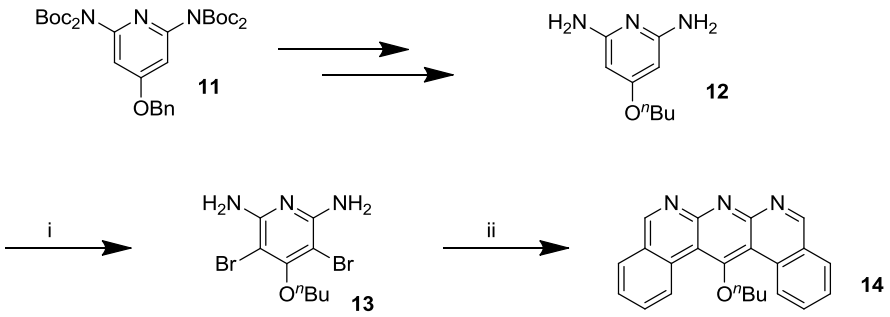
The synthetic routes to **6** and **7** are exceedingly simple and amenable to modification, being easy to scale up and taking only three total steps from the commercially available **8** (Scheme 4.1). The key step is a tandem Suzuki coupling and imine formation between halide **9** and boronic acid **10**, which forms the entire conjugated ring structure in one reaction.



Scheme 4.1 | Synthesis of triple HBA 7 (red) and triple HBD 6 (blue) from commercially available 2,6-diaminopyridine.[3]

Various strategies have been attempted to tether **7** *via* a variety of attachment points (Figure 4.8, indicated positions ***a-f***). Although the choice of attachment point would undoubtedly influence the geometry and precise properties of any resulting supramolecular polymer, synthetic accessibility is also a factor and a proof of principle was sought. Extensive efforts were made previously to tether **7** *via* points ***b-e***. These were carried out by S. Nakamura and B. Blight and are not documented here but were unsuccessful.

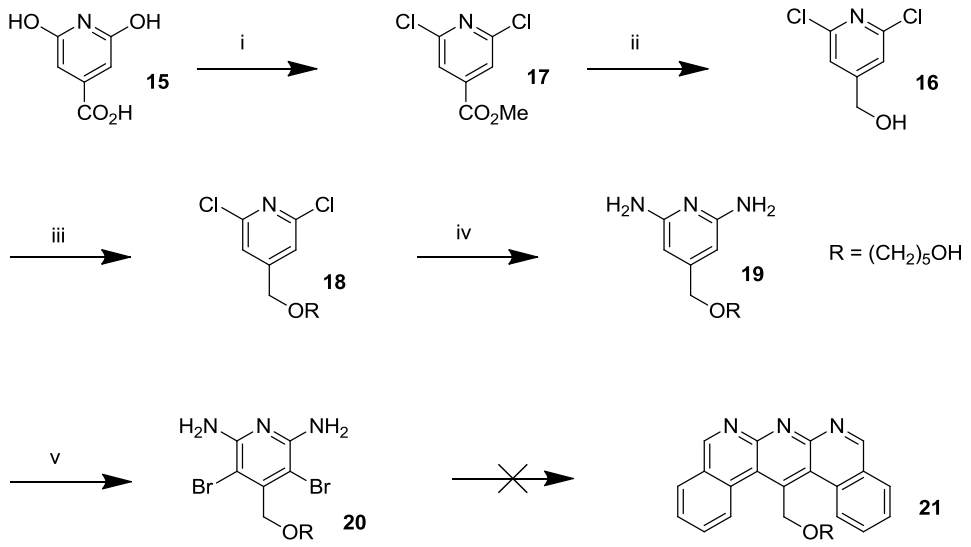
Tethering *via* position ***f*** was initially pursued. A hydroxypyridine motif had originally been developed (Scheme 4.2) based on the known compound **11**.[18] B. Blight provided a quantity of **12**, bearing an *n*-butyl ether as a model system, which could be halogenated with *N*-iodosuccinimide in good yield to give **13**. (Scheme 4.2)



Scheme 4.2 | A hydroxypyridine ether design was attempted for a tethered triple hydrogen bond acceptor motif 14, but failed at the final stage. Reagents and conditions: (i) NBS, DMF, -30°C then rt, 2h, 80% (b) (2-CHO)$C_6H_4B(OH)_2$, $Pd(PPh_3)_4$ DME, H_2O, Na_2CO_3, 1h, reflux, 25%

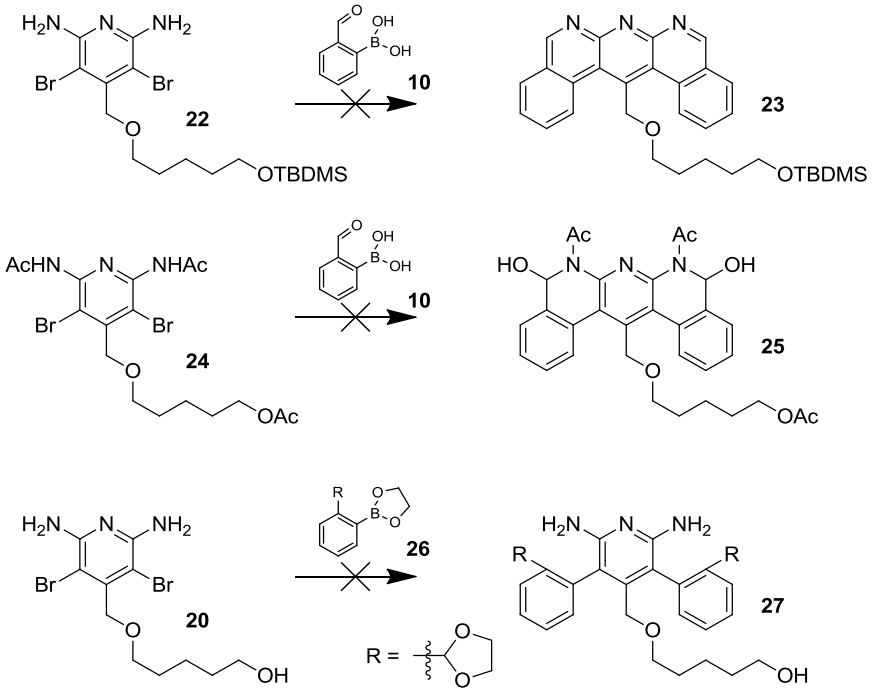
However, subsequent cyclisation with palladium and 2-formylphenylboronic acid only yielded a small quantity of **14** which could not be adequately purified, though the heterocycle had formed as evidenced by ^{1}H NMR, with the correct number of proton resonances in the range 7.5-9 ppm. This, coupled with the long, low-yielding and unscalable route to **12** from expensive starting materials (chelidamic acid at £10/g is transformed into **11** in 4 steps and 41% yield, with multiple stages of column chromatography) led to an alternative strategy being sought for any material suitable for use in a supramolecular polymer.

The cheap (£0.2/gram) starting material citrazinic acid **15** can be converted to the alcohol **16** using known chemistry: treatment with phosphorus oxychloride followed by a quench with methanol gave the methyl ester **17** directly, which could be readily reduced with sodium borohydride; the withdrawing effect of the pyridine ring enhanced the nucleophilicity compared to aliphatic or most aromatic esters. Alcohol **16** could be converted to the triflate (tosylation proved ineffective) and etherified with an excess of 1,5-pentanediol *in-situ* to give the alcohol-tethered chloride **18** in good yield. Conversion of **18** to an amine was not possible *via* a safer azide-displacement/reduction strategy, but treatment with ammonia and copper in a steel bomb gave the 2,6-diaminopyridine motif of **19**. Compound **19** could subsequently be halogenated to give the target molecule **20**. Column chromatography could be used to purify either **19** or **20**, but the overall yield for the steel bomb reaction and halogenation (~50% over two steps) remained the same regardless of whether crude or pure **19** was used in the iodination step. (Scheme 4.3)



Scheme 4.3 | A benzyl ether design was attempted for a tethered triple hydrogen bond acceptor motif 21, but failed at the final stage. Reagents and conditions: (i) $POCl_3$, NMe_4Cl, reflux, 16h, 83% (ii) $NaBH_4$, EtOH, rt, 16h, 92% (iii)TEA, Tf_2O then $CH_2(CH_2CH_2OH)_2$, 94% (iv) NH_3 (aq), Cu, 180 °C, 16h, 80% (v) NIS, DMF, -30 °C then rt, 2h, 60%

Unfortunately, the final step to form the tethered acceptor **21** was not successful, despite several modifications (Scheme 4.4).



Scheme 4.4 | Attempts to change the substrates for the final cyclisation step towards 21.

Initially, **20** was protected as the TBDMS ether **22** to try and overcome any potential solubility issues or involvement of the free hydroxyl group of **20**, however no **23** could be detected from cross-coupling reactions. Acylation of **20** to **24** would also protect the alcohol, as well as the amino groups; the acyl substituents could then be removed from **25** to perform the final aromatisation step, but again the final cross-coupling was never successful. Next, the aldehyde group of **10** was protected as acetal **26**, and it was envisioned that **27** could be deprotected and cyclised after the key C-C bond forming step, but again the final cross-coupling was never successful.

The only remaining site available was position ***a***, previously overlooked due to anticipated problems with steric bulk in the vicinity of the hydrogen bonding face. However, a model experiment (Figure 4.9) indicated that binding between **28** and **29** was possible given the required substitution pattern, possibly due to a slightly staggered binding conformation similar to that seen in the crystal structure of **6•7** (Chapter I, Figure 1.13).[17]

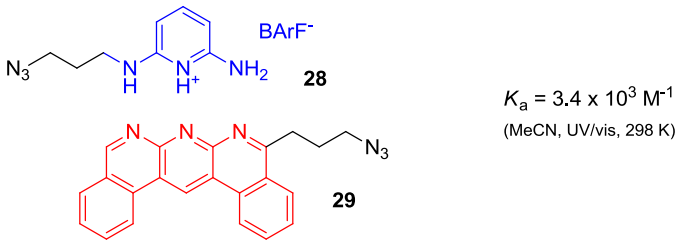
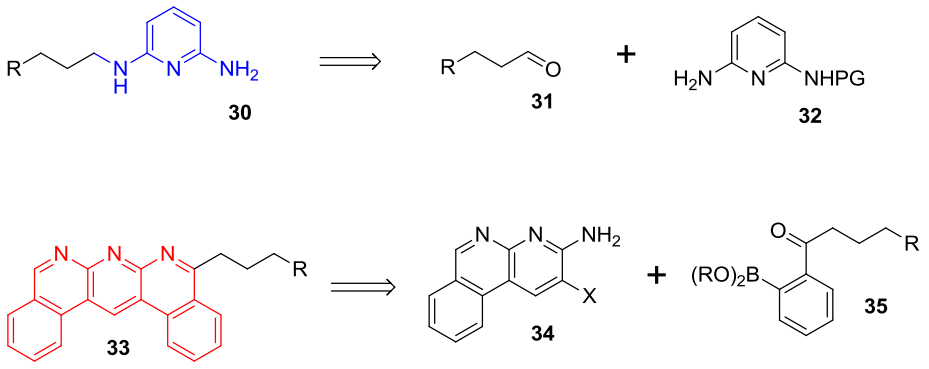


Figure 4.9 | Binding of Acceptor (red) and Donor (blue) is largely unaffected by substitution in positions adjacent to the recognition motif.

In order to access mono-substituted versions of AAA **7** and DDD **6**, a way was needed to make unsymmetrical versions of synthetic intermediates, and the proposed retrosynthesis is shown in Scheme 4.4.

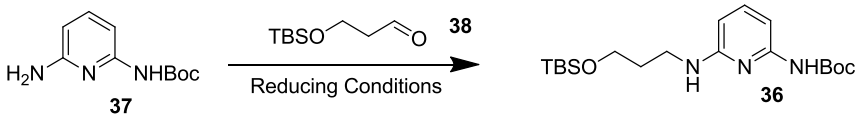


Scheme 4.5 | A retrosynthesis of unsymmetrical donor (top) and acceptor (bottom) motifs.

A tethered donor **30** could be accessed by a reductive amination of a suitably functionalised aldehyde **31** and an unsymmetrical, mono-protected diaminopyridine such as **32**. A tethered acceptor **33** could be synthesised from known monofunctionalised naphthyridines such as **34**, and an acyl-substituted boronic acid such as **35**. In both cases, the R groups should be a functionality that can be easily and selectively elaborated, such as an alcohol or azide.

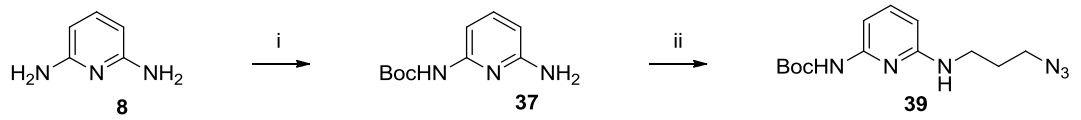
4.3.2 Synthesis of tethered donor motif.

In order for a hydrogen bond donor motif to be tethered in a convergent synthetic strategy, the tether must have some functionality which can be elaborated on using chemistry orthogonal to the functionality of the acceptor motif itself. Alcohols and amines could be elaborated using electrophiles and would require protecting groups or selective chemistry in order to effect discrimination. An alcohol tethered donor **36** was initially synthesised by reductive amination of aminopyridine **37** and aldehyde **38**, by Dr Barney Walker (Scheme 4.5), but suffered from poor versatility.



Scheme 4.5 | The key step of a successful synthesis of an alcohol-tethered donor motif, by Dr Barney Walker.

Pericyclic chemistry has the potential to be selective, high-yielding and mild, and of particular versatility is the copper-catalysed alkyne-azide cycloaddition (CuAAC, or "click" reaction). Azide groups in particular are non-polar and stable to most reaction conditions, and can usually be introduced early in a synthetic sequence without deleterious effects. The other partner in the CuAAC is an alkyne, which usually require protecting groups and so azides were chosen to elaborate the acceptor motif (Scheme 4.6).



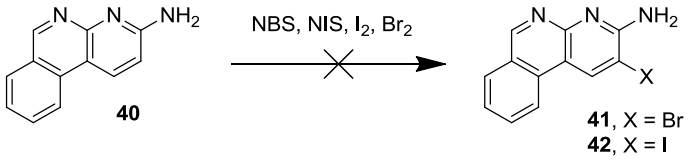
Scheme 4.6 | Synthesis of an azide-tethered donor motif 39. Reagents and conditions: i) Boc_2O, THF, 60 °C, 16 hours, 63% ii) $N_3CH_2CH_2CH_2OH$ with DMP, DCM, rt, then amine and $NaBH(OAc)_3$, rt, 90 min, 68%.

The desymmetrised **37** was obtained in reasonable yield *via* a Boc protection of **8**. The aldehyde for the corresponding reductive amination is 3-azidopropanal, which was so volatile that isolating the aldehyde from oxidation of 3-azidopropanol was never successfully carried out.* Thus, a one-pot procedure was developed where the known[19] 3-azidopropanol was oxidised with Dess-Martin Periodinane (DMP). Without any workup or treatment, amine **37** was added and acetic acid from the reduction of DMP catalysed the formation of the corresponding imine, and sodium triacetoxyborohydride selectively reduced the imine to **39**. After evaporation and chromatographic purification, **39** was isolated in 68% yield and a 90 minute reaction time from a 1-pot, 3-reaction procedure.

*As a short-chain azide with 41% by mass of nitrogen atoms, isolation was also considered an avoidable hazard.

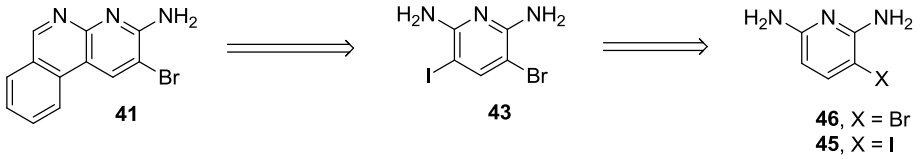
4.3.3 Synthesis of tethered acceptor motif.

The tethered acceptor was anticipated to be accessible *via* a heterocyclic core with three of the five rings already in place. Although in principle aminonaphthyridine **40** could be electrophilically halogenated to the bromide **41** or iodide **42** to provide the desired functionality, in practice the selectivity was compromised by the reduced reactivity compared to diaminopyridine (Scheme 4.7) and this approach was never successful, leading to decomposition or unidentifiable insoluble products.



Scheme 4.7 | The first unsuccessful approach to an unsymmetrical triply hydrogen bonded acceptor precursor.

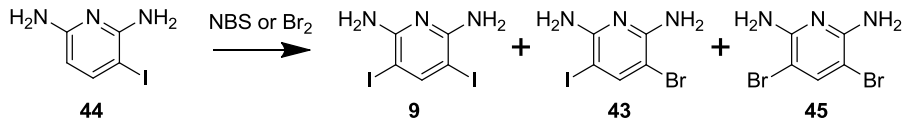
An alternative approach was envisaged whereby a differentially dihalogenated diaminopyridine **43** could be selectively mono-reacted with an appropriate boronic acid, providing rapid access to multigram amounts of **41**. (Scheme 4.8).



Scheme 4.8 | Synthetic pathway to access unsymmetrical triple hydrogen bond acceptors.

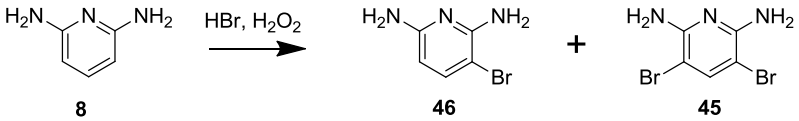
The first step would be accessing highly pure **43**. Unfortunately, attempts to brominate iodide **44** always lead to some scrambling (Scheme 4.9), as *N*-bromosuccinimide can efficiently perform *ipso*-substitution on the site of the iodine and generate small amounts of *N*-iodosuccinimide *in-situ*, further reacting to generate all three possible products in significant amounts. These could not be efficiently

purified on a large scale, and contamination with diiodide **9** and dibromide **45** lead to major purification problems downstream.



Scheme 4.9 | Bromination of iodide 44 lead to scrambling, typically 9:43:45 of 1:5:2.

Reversing the order of halogenation reactions, and carrying out an iodination on the known 3-Bromo-2,6-diaminopyridine **46** was much more successful, completely preventing scrambling. However, the purity of **46** was very important, as the reaction is not completely selective and the product contains between 5-20% of the over-bromination product **45** as well.

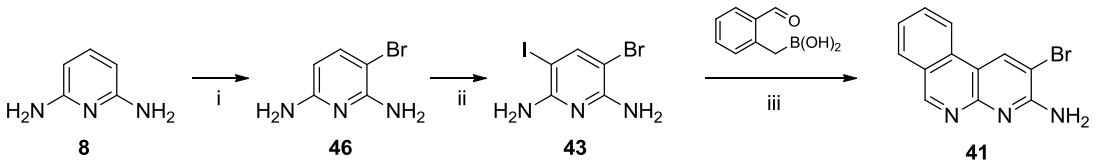


Scheme 4.10 | Synthesis of 46 was not completely selective, typical ratio 46:45 in the crude reaction mixture was 85:15. After recrystallisation from EtOAc and toluene, typically > 99:1.

After considerable experimentation, optimal recrystallisation conditions (EtOAc then toluene) were found to furnish 99% pure **46** in 100 gram quantities (Scheme 4.10). The differing solubility of **46** and **43** could be exploited in the next step to give an ambient and mild iodination with NIS in DMF/MeOH that precipitates **43** in a consistent 61-64% yield and 99% purity (4 batches, 2 – 16 grams).

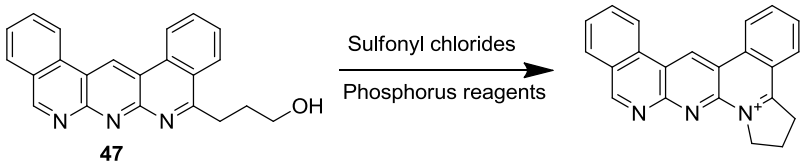
For the first Suzuki coupling, contamination of **41** with unreacted **43** was found to be far more difficult to purify in a subsequent step than when contaminated with the doubly-cyclised **7**, so 1.5 equivalents of the boronic acid were used. The actual purity of **41** approaches 90% due to competing protodeborylation of the boronic acid to benzaldehyde. Up to this point all products were isolated by precipitation and purified

by recrystallisation, meaning that 10+ gram batches of **41** could be prepared very rapidly. (Scheme 4.11)



Scheme 4.11 | Synthesis of 41. Reagents and conditions: i) 48% HBr, H_2O_2, rt, 30 min, 74% then recrystallisation from EtOAc (70%) and toluene (99%). ii) NIS, MeOH:DMF 2.5:1, rt, 30 min, 64% iii) (2-CHO)C_6H_4B(OH)$_2$, Pd(PPh$_3$)$_4$, K_2CO_3, H_2O:1,4-Dioxan 1:3, 80°C, 16h, 88%, with 10% of 7.

The tether functionality itself was initially introduced as an alcohol by Dr. Barney Walker, using similar chemistry to that in Scheme 4.13, but this gave some rather perplexing problems when conversion of the alcohol **47** to azide **29** was attempted, normally a trivial and reliable reaction. It was thought that conversion of the alcohol of **47** to a leaving group lead to intramolecular reactions that either out-competed the azide anion or did not survive workup (Scheme 4.12).

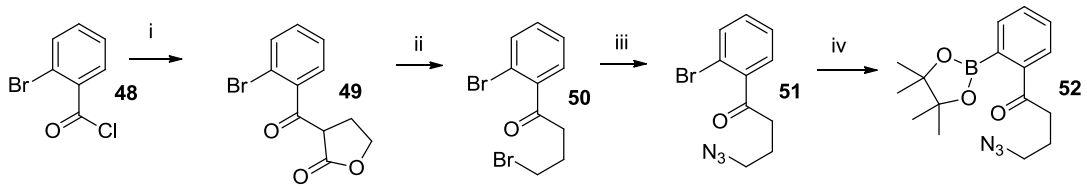


Scheme 4.12 | Problems with late-stage azide incorporation, explored by Dr. Barney Walker.

Since methods for conversion of an alkyl alcohol to an azide usually entail a nucleophilic intermediate, an alternative strategy was sought. Azide groups are inert to a variety of reaction conditions, so one was introduced earlier in the synthesis *via* a substituted boronic acid.

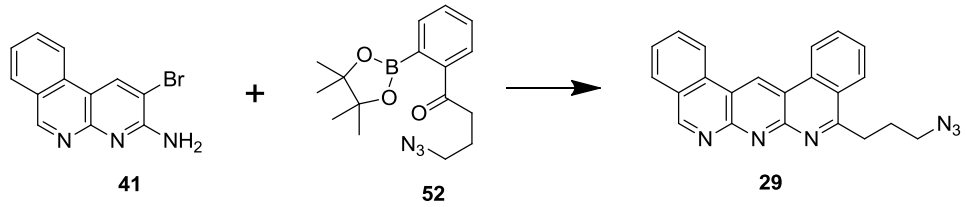
2-Bromobenzoyl chloride **48** can be reacted with the enolate of γ-butyrolactone to furnish a substituted lactone **49**, which can be be ring-opened, decarboxylated, and brominated by heating with HBr to give bromide **50**. Compound **50** can then be subject to an azide displacement in DMF, giving **51** in a sequence requiring only three

workups and then a single filtration through a plug of silica in 69% overall yield and decagram scale. Compound **51** can be converted to boronate ester **52** *via* a borylation with bis(pinacol)diborane, avoiding the need for ketone protecting groups or strong bases (Scheme 4.13). Triphenylphosphine is one of the few reagents that reacts readily with azides, so the use of palladium catalysts with triphenylphosphine ligands were avoided.



Scheme 4.13 | Synthesis of tethered boronate 52. Reagents and conditions: i) γ-butyrolactone, LDA, THF, -78°C, 1 hour ii) 48% HBr, 80°C, 20 min iii) NaN_3, DMF, 80°C, 30 min, 75% over 3 steps iv) B_2Pin_2, DMSO, KOAc, $Pd(dppf)Cl_2$•CH_2Cl_2, 80°C, 30 min, 97%

The final target was then assembled with another Suzuki coupling between **41** and **52**, in low yield due to losses in workup and purification, but able to furnish gram quantities of pure **29** which was sufficient for all further studies. (Scheme 4.14).

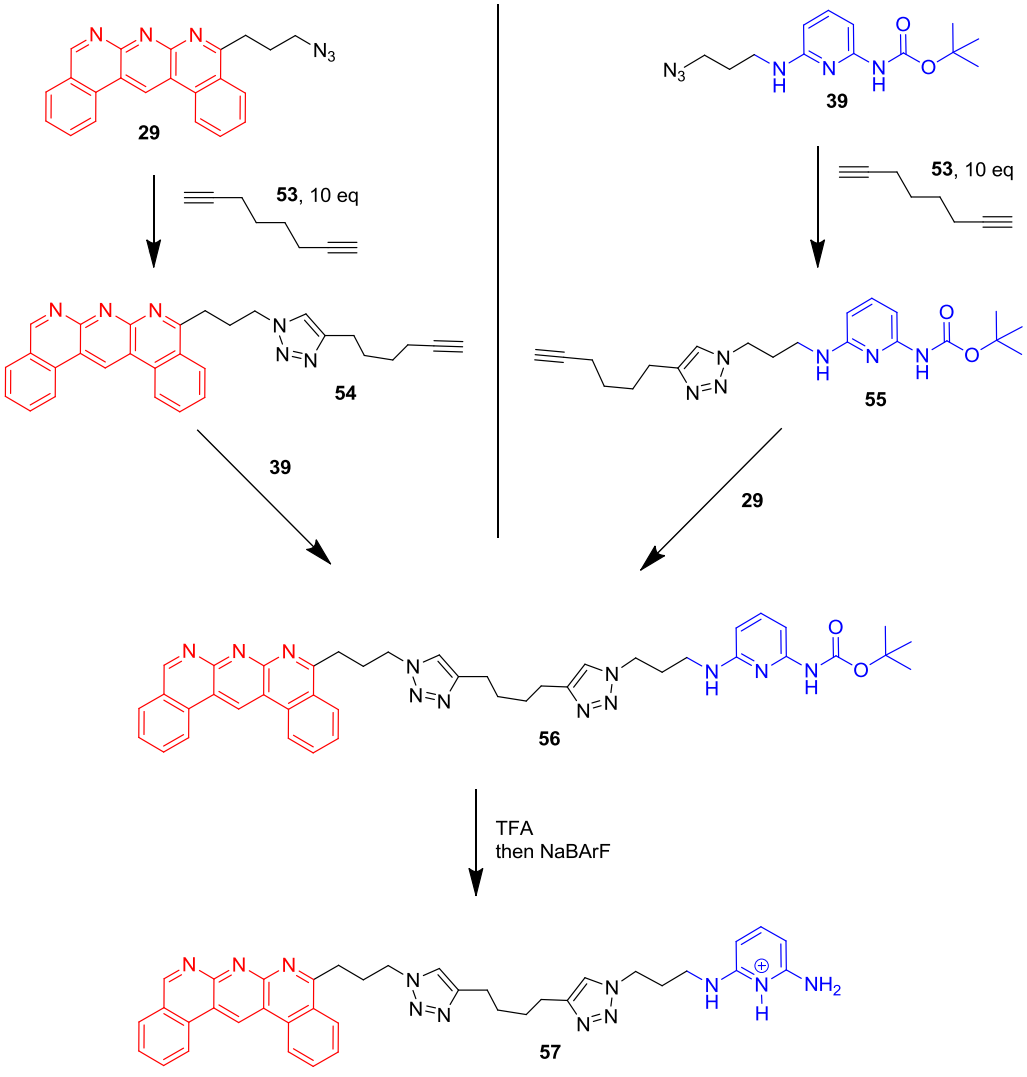


Scheme 4.14 | Synthesis of tethered acceptor. Reagents and conditions: K_2CO_3, $Pd(dppf)Cl_2$•CH_2Cl_2, H_2O:1,4-Dioxan 1:3, 80°C, 45 min, 29%

4.4 Development of a polymer linker

4.4.1 Simple alkyl linkers

In order to make best use of building blocks **29** and **39**, a commercial and inexpensive dialkyne **53** was chosen as the linker.



Scheme 4.15 | Attempted stepwise assembly of a supramolecular polymer precursor.

By using an excess of **53**, it can be selectively mono-substituted with either an AAA unit **29** or a DDD precursor **39** to give **54** or **55** respectively. These can be clicked

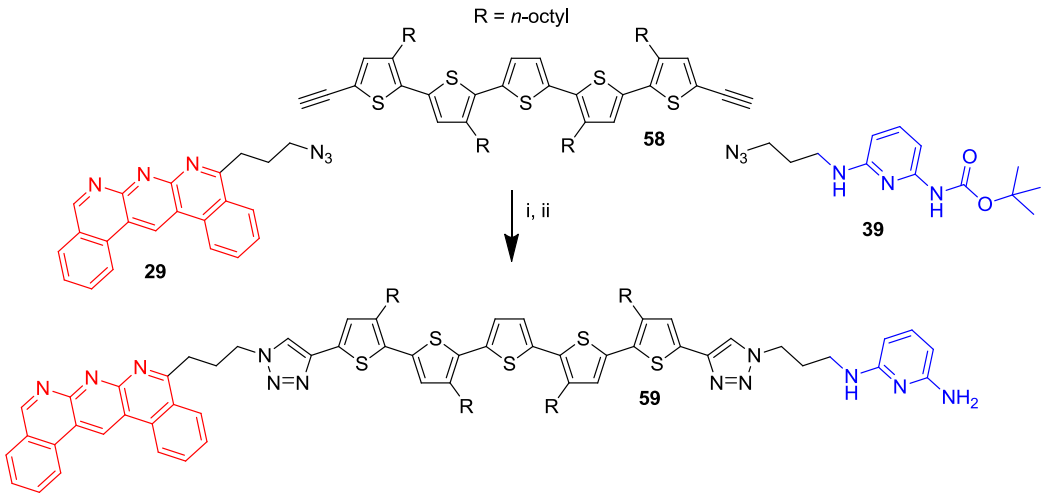
together with the respective complementary building blocks to assemble polymer precursor **57**, with either route giving yields of ~50% over two steps. (Scheme 4.15).

It was anticipated that molecule **56** could be treated with acid to cleave the Boc protecting group, then be subjected to ion exchange to provide the target polymer directly. However, when deprotonated during workup with an aqueous solution of sodium carbonate, the neutral form of **57** proved to be insoluble in any solvents that would not disrupt the formation of the hydrogen bonding network, and so an alternative solubilising linker was sought.

4.4.1 Functional linkers

Oligo- and poly-thiophenes are being widely used as conductive organic materials,[20,21] and a pH-controlled way of reversibly self-assembling short chains of thiophenes would be highly desirable. To this end, pentathiophene **58** was provided by Mr M. Gall and Dr Barney Walker.[22]

The linker was installed in a one-pot, three-component reaction that yielded the desired heterodimer along with two side products. Longer strategies could be envisioned that gave more efficient use of building blocks **29** and **39**, but sufficient material was generated for initial studies this way. Because the AAA motif from buiding block **29** is relatively polar, the product could be separated by standard column chromatography to give the desired heterodimer in 24% yield. The Boc group could then be removed under standard conditions (DCM:TFA 1:1), followed by a stoichiometric protonation with HCl and salt exchange with $NaBArF^-$ to give the target molecule with a solubilising counterion. The $BArF^-$ salt was not isolated, but made as needed and used *in-situ*, referred to as $[\mathbf{59}+H]^+$.



Scheme 4.16 | Synthesis of 59. Reagents and conditions: i) $Cu(MeCN)_2 \bullet PF_6$, MeOH:DCM:TEA 10:2:1, 24% ii) TFA:DCM 1:1, 64%. A synthesis of 58 is given in section 4.7 as carried out by Malcolm Gall.

4.5 Characterisation of a supramolecular polymer

4.5.1 In-house NMR analysis

The characterisation of supramolecular polymers is hampered by the dynamic nature; most analytical techniques for traditional polymers (gel permeation chromatography, mass spectroscopy) entail a change in solvent, conditions, concentration or temperature. A conventional polymer has a static chain length, but supramolecular polymers are exceedingly sensitive to these changes due to the dynamic nature of the intra-monomer interactions. However, the first step was to establish that the monomers were self-associating in some way.

By proton NMR, self-association can be seen when **59** is "activated" by protonation (spectrum "2" to spectrum "1"). For comparison, the spectrum of compound 16 is shown in the absence (spectrum "3") and presence (spectrum "4") of an appropriate binding partner, showing the same characteristic shift of the **A**/**a** protons compared to **F**/**f**, taken to be indicative of a recognition event.

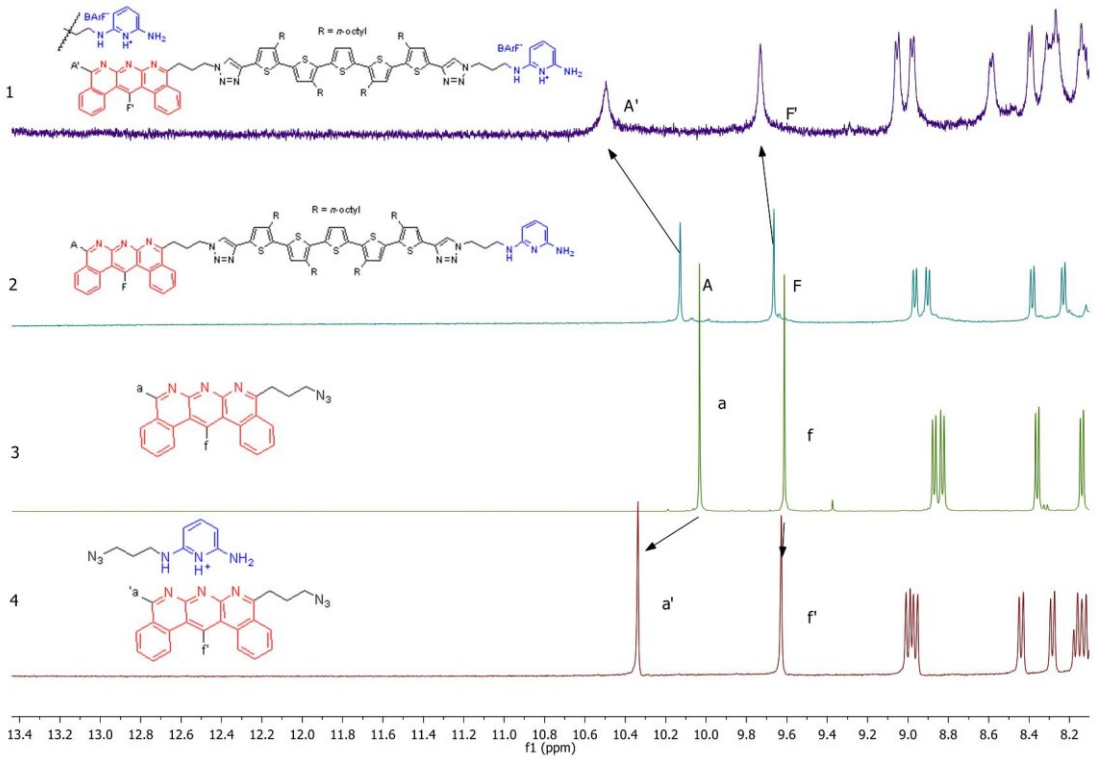


Figure 4.10 | 1) 59 plus 1 eq acid 2) 59 3) 29 4) 29 plus 1 eq 28. Omitted for clarity: The rightmost blue "donor" motif in 1 is complexed to another red "acceptor" motif.

When complexation happens, there is a pronounced (~0.3 ppm) deshielding effect on the "ortho" protons **A** and **a** respectively, thought to be due to their proximity to the recognition site. However, the "para" proton in position **F** and **f** is only marginally altered, by inductive rather than steric effects. This is a handy "fingerprint" for detecting the formation of a triply hydrogen bonded complex, and is not observed when the acceptor motif simply protonates (Chapter III, Figure 3.15). additionally, the recognition event was largely independent of concentration. With a K_a in non-polar solvents ($CDCl_3$) on the order of 10^{10} M^{-1}, an AAA•DDD would exist only in the complexed state until a concentration of approximately 10^{-10} M^{-1}, five orders of magnitude below the detection limit for NMR spectroscopy.

Having established that the new system is undergoing switchable self-recognition, the remaining question is that of chain length: under what conditions will **59** self-assemble to form macrocyclic versus polymers? Peak broadening across the whole of the ^{1}H NMR spectrum may indicate aggregation into large species, and is only evident in the cationic ditopic monomer (spectrum 1, Figure 4.10) and not when the

unfunctionalised monomers bind together (spectrum 4, Figure 4.10). However, a more rigorous analysis was sought.

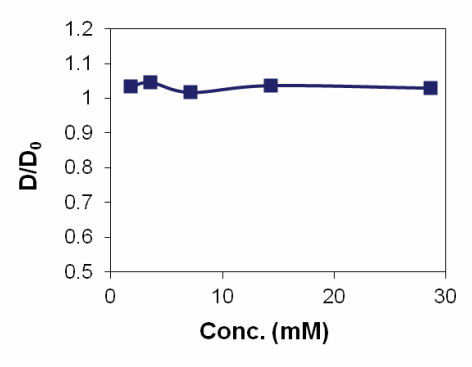


Figure 4.11 | Normalised DOSY diffusion coefficient of a solution of [59+H]$^+$ (CD_2Cl_2, 298 K) at various concentrations.

DOSY (diffusion-ordered spectroscopy) is a 2-dimensional NMR technique that gives a relative readout of diffusion speed (related to molecular size by the Einstein-Stokes equation). Thus, a species that changes size (e.g. number of monomer units) with concentration ought to be easily characterised by a series of DOSY experiments. Unfortunately the results seemed to indicate that the material was not forming polymers, showing a poor relationship between the relative diffusion coefficient (standardised to solvent residual peaks) and concentration (Figure 4.11). However, the 1-dimensional ^{1}H NMR spectrum, which was concomitantly measured, had some interesting concentration-dependent features around 3 mM which may be indicative of aggregation (Figure 4.12).

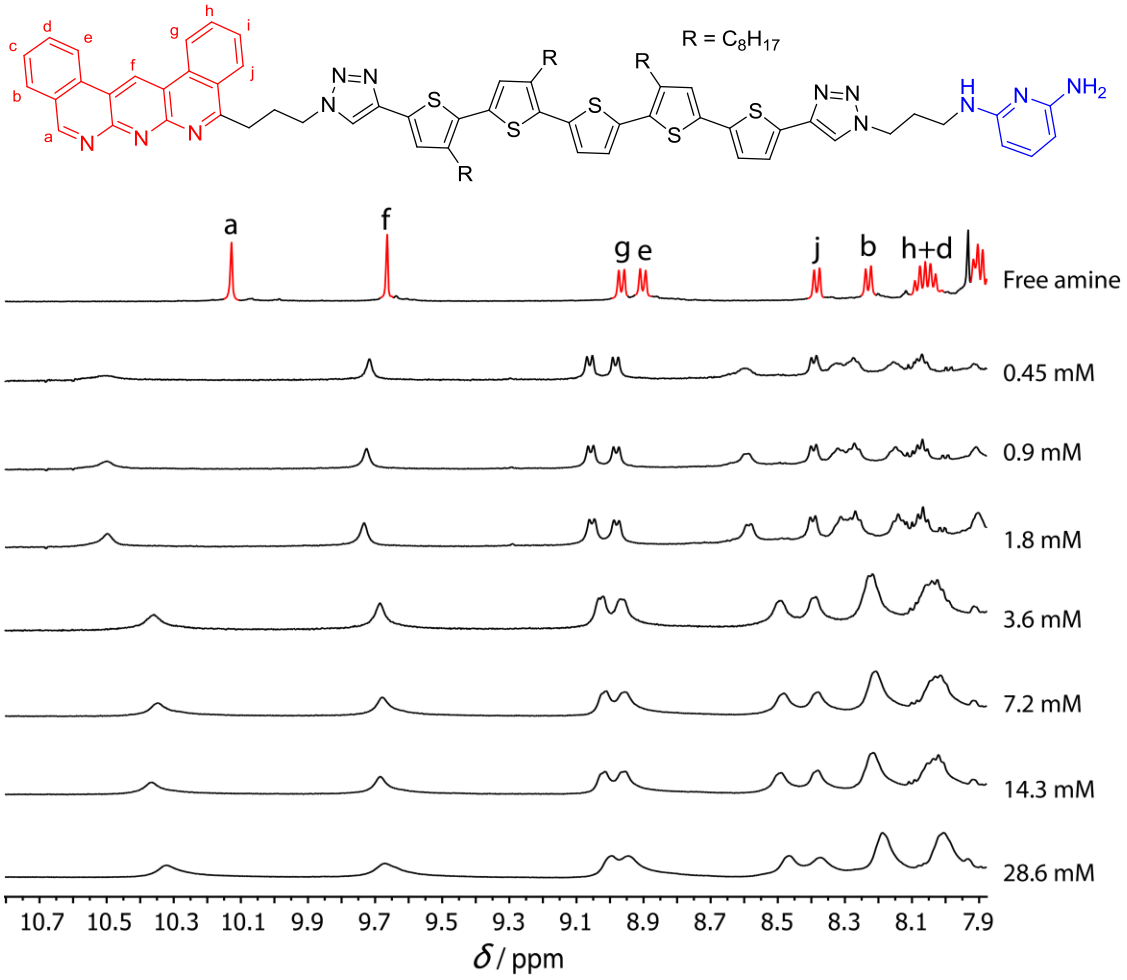


Figure 4.12 | 1) [59+H]$^+$ at various concentrations, with the spectrum of the unprotonated (and therefore monomeric) 59 for comparison, top.

4.5.2 Collaborative NMR analysis

In collaboration with the research group of Prof. Rint Sijbesma (Eindhoven), a small sample of **59** was analysed more rigorously by NMR. The diffusion coefficients of [**59**+H]$^+$ were plotted but only show a 40% increase in diffusion coefficient over the concentration range of the experiment (Figure 4.13). Unfortunately, this seems to indicate that the material probably exists as small macrocycles, as a purely linear supramolecular polymer would experience a three-fold increase in diffusion coefficient over the same range. This could be due to the rigidity of the thiophene backbone actually promoting macrocycle formation, as discovered with less conformationally-flexible linkers by Meijer.[3] Even in Meijer's more flexible linkers, polymeric species were only observed above 75 mM, whereas **59** suffered from solubility problems due to thiophene π-stacking.

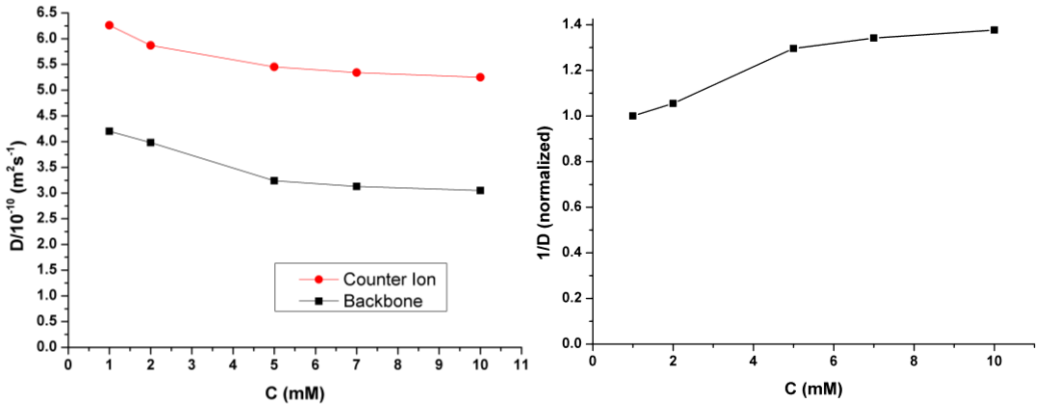


Figure 4.13 | Left, concentration dependence of the diffusion coefficient of 59, inversely related to molecular volume. Right, a normalised plot of 1/D showing a 40% increase in diffusion coefficient. Results and plot thanks to Abidin Balan and Marko Nieuwenhuizen (Sijbesma group)

4.5.3 Collaborative materials and AFM analysis

Samples of **59** and [**59**+H]$^+$ were sent to the laboratory of Prof. Tim Swager (MIT) where Dr Barney Walker incorporated the material into organic solar cells and organic field effect transistors. Unfortunately, results were negative, indicating that the charges present on **59** as a necessary consequence of hydrogen bond motif formation also preclude its use in organic electronic materials. Additionally, both materials were analysed by atomic force microscopy by Dr Barney Walker for signs of formation of fibres that could indicate polymer self-assembly, but these results were inconclusive.

4.6 Conclusions and future work

A major aim of the current project was to develop a synthetically accessible, scalable, versatile synthon for acid/base-switchable hydrogen bonding motifs. AAA and DDD motifs were tethered in a way that compromised neither the binding nor the switching, and left the versatile and flexible azide group available for elaboration *via* well-understood CuAAC "click" chemistry. This has been achieved, and an initial attempt to elaborate the design with a complex thiophene-based linker was synthetically successful but did not give the desired functionality due to conformational and solubility effects.

The focus of the project in the future will be uses of this motif. A highly desirable target is a simple ditopic AAA-DDD monomer, with a linker that provides both conformational flexibility and solubility since supramolecular polymers are often characterised at high concentration (> 100 mM). Fortunately, solubilising groups can be incorporated without altering the synthesis or structure of the recognition motifs, and we envision the incorporation of easily-accessible and solubilising phenylene spacers such as in **60** (Figure 4.14).

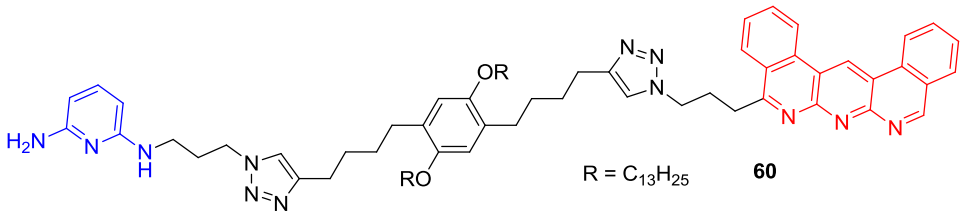


Figure 4.14 | Proposed structure of a more soluble AAA-DDD dimer.

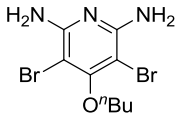
Another possibility is the use of these hydrogen bonding motifs to make crosslinking gels which are acid-base responsive, similar to Zimmerman's redox-switching gels (Figure 4.7).[15]

4.7 Experimental procedures

4.7.1 Synthesis of new compounds

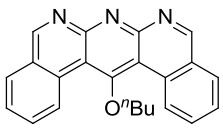
Unless stated otherwise, all reagents and solvents were purchased from Aldrich Chemicals and used without further purification. Column chromatography was carried out using Silica 60A (particle size 35-70 μm, Fisher, UK) as the stationary phase. Dry Column Vaccum Chromatography was carried out according to the procedure of Pedersen[23], using silica gel 60 (particle size 15-40 μm, Merck, UK) and TLC was performed on precoated silica gel plates (0.25 mm thick, 60 F254, Merck, Germany) and observed under UV light. NMR spectra were recorded on Bruker AV 400, and Bruker DMX 500 instruments. Chemical shifts are reported in parts per million (ppm) from low to high frequency and referenced to the residual solvent resonance. Coupling constants (*J*) are reported in hertz (Hz). Standard abbreviations indicating multiplicity were used as follows: s = singlet, d = doublet, t = triplet, dd = double doublet, q = quartet, m = multiplet, b = broad, ddd = doublet of double doublets. ^{1}H and ^{13}C NMR assignments were made using 2D-NMR methods (COSY, ROESY, TOCSY, HSQC, HMBC) and are unambiguous unless stated otherwise. Melting points (m.p.) were determined using a Sanyo Gallenkamp apparatus and are reported uncorrected. Low resolution ESI mass spectrometry was performed with a Finnigan LCQ-MS, Micromass Platform II or Waters Quattro Ultima LC-*MS*/*MS* mass spectrometers. High resolution ESI and FAB mass spectrometry were carried out by the mass spectrometry services at the University of Edinburgh and the EPSRC National Mass Spectrometry Service Centre, Swansea, UK. The following compounds were prepared as indicated: **7** and **6** were prepared as reported previously.[17] Binding constants were determined by UV/vis spectroscopy as reported previously (Chapter II, Section 2.7.2). Some compounds had incomplete data available and were included for general interest but would require re-analysis if remade: **13**, **14**, **16-20**. Mr Malcolm Gall synthesised and characterised compounds **61-65**. Dr Barney Walker developed the synthetic protocol for compounds **66**, **59**, and **39** and synthesised compounds **66** and **59**.

4-Butoxy-3,5-diiodopyridine-2,6-diamine (13)



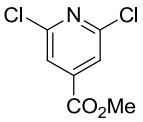
4-Butoxypyridine-2,6-diamine **12** (50 mg, 0.276 mmol) was provided by Dr B. Blight, and was dissolved in DMF (1 mL). To this was added a solution of *N*-bromosuccinimide (137 mg, 0.607 mmol) in DMF (1 mL) over 1 hour at -30 °C under a dry nitrogen atmosphere. The reaction mixture was warmed to room temperature and stirred for a further hour, then poured into aqueous Na_2CO_3 (2 M, 10 mL) and extracted with ethyl acetate (2 × 10 mL). The combined organic washings were filtered through a short plug of silica, which was then flushed with methanol (10 mL). The combined organic filtrates were evaporated *in vacuo* to give **13** as a brown tacky solid (97 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.80 (br.s, 4H), 3.89 (m, 2H), 1.82 (m, 2H), 1.51 (m, 2H), 0.94 (t, 3H).

14-Butoxybenzo[*f*]isoquino[3,4-b][1,8]naphthyridine (14)



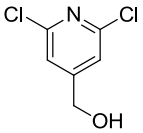
2-Formylphenylboronic acid (61.8 mg, 0.42 mmol) was dissolved in 1,2-dimethoxyethane (3 mL) and aqueous Na_2CO_3 (1.5 M, 2 mL). To this was added **13** (60 mg, 0.138 mmol) and $Pd(PPh_3)_4$ (15.9 mg, 0.013 mmol). The mixture was heated to reflux for 1 hour under a dry nitrogen atmosphere, then cooled and taken up in ethyl acetate (10 mL) and water (10 mL) the organic layer was separated, dried over magnesium sulfate, and concentrated *in vacuo*, then purified by dry column vacuum chromatography (10% methanol, 90% chloroform) to give **14** as an impure dark solid (12 mg, 25%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 9.12 (s, 2H), 8.82 (d, *J* = 7.7 Hz, 2H), 7.84 (d, *J* = 7.7 Hz, 2H), 7.58 (t, *J* = 7.7 Hz, 2H), 7.43 (t, *J* = 7.7 Hz, 2H). Resonances corresponding to the O^{n}Bu group were unclear due to masking by significant impurities.

Methyl 2,6-dichloropyridine-4-carboxylate (17)



Citrazinic acid **15** (100 g, 645 mmol) was dissolved in phosphorus oxychloride (180 mL, 1.92 moles, 3 equivalents) and to this was added tetramethylammonium chloride (77 g, 705 mmol). The reaction mixture was heated at 130 °C under reflux for 16 hours under a dry nitrogen atmosphere, then cooled and slowly poured into cold methanol (1 L). This solution was stirred for 30 minutes, then diluted with water (1 L) and neutralised with solid sodium bicarbonate (540 g, 6.4 moles). The aqueous was extracted with ethyl acetate (2 × 2 L), dried over magnesium sulfate, and filtered through a plug of silica, then concentrated *in vacuo* to give **17** as an earthy brown solid (107.8 g, 83%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.74 (s, 2H), 3.91 (s, 3H) in good agreement with literature data.[26]

(2,6-Dichloropyridin-4-yl)methanol (16)

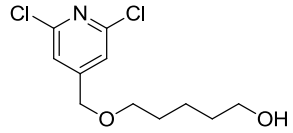


Compound **17** (106.77 g, 526 mmol) was dissolved in ethanol (500 mL) and cooled to 0 °C. To this was added portionwise powdered sodium borohydride (80 g, 2.11 moles, 4 equivalents), with the addition rate being sufficient to heat the mixture to a gentle reflux.* The reaction mixture was then stirred overnight at room temperature, and quenched with aqueous HCl (1 M, 1 L). The mixture was then made alkaline with a minimum volume of sodium carbonate (aq., saturated) and extracted with dichloromethane (2 × 1 L). The combined organic layers were dried over magnesium sulfate and concentrated *in vacuo* to give **16** as a tan solid (86.3 g, 92%). ^{1}H NMR

* This was not intended experimental protocol, and the temperature should be kept at 0 °C throughout the addition in future, heating to reflux only after an extended period at room temperature, and then only if conversion is incomplete by TLC.

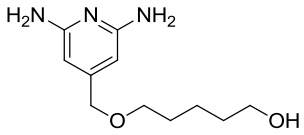
(400 MHz, $CDCl_3$): δ = 7.21 (s, 2H), 4.68 (s, 2H) in good agreement with literature data.[26]

5-[(2,6-Dichloropyridin-4-yl)methoxy]pentan-1-ol (18)



Compound **16** (5.93 g, 33.3 mmol) and triethylamine (3.36 g, 4.63 mL, 33.3 mmol), was dissolved in dry dichloromethane (80 mL) and this was added dropwise to a solution of triflic anhydride (9.39 g, 33.3 mmol) in dry dichloromethane (40 mL) over 30 minutes at 5 °C under a dry nitrogen atmosphere. The solution was stirred at room temperature for a further hour and cooled to 0°C, then 1,5-pentanediol (30 mL, 290 mmol, 9 eq.) was added and the reaction mixture stirred vigorously at room temperature for a further hour. The reaction mixture was then diluted with ethyl acetate (360 mL) and washed with HCl (1 M, 2 × 500 mL) and sodium bicarbonate (saturated aq., 2 × 500 mL). The organic layer was dried over magnesium sulfate and concentrated *in vacuo* to give **18** as a yellow oil (8.26 g, 94%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.16 (s, 2H), 4.41 (s, 2H), 4.05 (m, 1H), 3.61 (m, 2H), 3.46 (m, 2H), 1-64-1.41 (M, 4H), 1.19 (m, 2H).

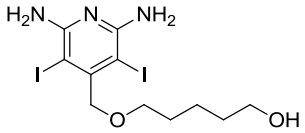
5-[(2,6-Diaminopyridin-4-yl)methoxy]pentan-1-ol (19)



Compound **18** (2.80 g, 10.6 mmol) was dissolved in saturated aqueous ammonia (32 mL) and to this was added copper powder (211 mg, 3.32 mmol). The mixture was placed in a steel bomb and heated to 180 °C for 16 hours, with appropriate safety shielding used. The reaction was cooled to room temperature, then the safety shield removed and the copper removed by filtration. The filter was washed with further water (50 mL), then extracted with a 3:1 mixture of chloroform and isopropyl alcohol (4 × 100 mL). The combined organic layers were dried over magnesium sulfate, then concentrated *in vacuo* to give **19** as a green solid (1.85 g, 80%). A portion was

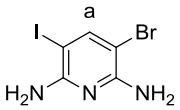
purified by dry column vacuum chromatography (10% methanol, 90% chloroform) for further analysis. ^{1}H NMR (400 MHz, DMSO): δ = 5.52 (s, 2H), 5.29 (br.s, 4H), 4.30 (br.s, 1H), 4.09 (s. 2H), 3.30 (m, 4H), 1.48-1.21 (m, 6H). ^{13}C (126 MHz, DMSO): δ = 158.6, 149.8, 93.3, 71.1, 69.7, 60.6, 32.3, 29.1, 22.2.

5-[(2,6-Diamino-3,5-diiodopyridin-4-yl)methoxy]pentan-1-ol (20)



Compound **19** (450 mg, 2 mmol) was dissolved in DMF (8 mL). To this was added a solution of *N*-iodosuccinimide (922 mg, 4.1 mmol) in DMF (8 mL) over 1 hour at -30 °C under a dry nitrogen atmosphere. The reaction mixture was warmed to room temperature and stirred for two hours, then diluted with ethyl acetate (50 mL) and washed with lithium chloride (5% w/v aq., 4 × 50 mL). The organic layer was dried over magnesium sulfate and evaporated *in vacuo* to give **20** as a brown solid (580 mg, 60%). ^{1}H NMR (400 MHz, DMSO): δ = 5.73 (br.s, 4H), 4.49 (s, 2H) 4.28 (br.s, 1H), 3.39 (m, 2H), 3.31 (m, 2H), 1.46-1.24 (m, 6H).

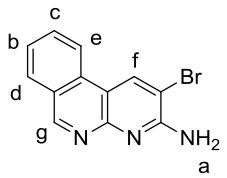
3-Bromo-5-iodopyridine-2,6-diamine (43)



3-Bromopyridine-2,6-diamine **46** was prepared according to the procedure of Boudakin[24] and recrystallised from EtOAc (10 ml/g, 70%) and PhMe (15 ml/g, 99%) before use. To a solution of 3-bromopyridine-2,6-diamine (15.04 g, 80 mmol) in DMF (160 ml) and MeOH (400 ml) was added a solution of *N*-iodosuccinimide (17.1 g, 0.95 eq) in DMF (160 ml) dropwise with stirring over 30 minutes at room temperature. The reaction mixture was stirred for a further 30 minutes at room temperature. The precipitate was collected by filtration and washed with methanol, then dried to give **43** as a grey powder (16.01 g, 64%), mp 219 °C (MeOH/DMF,

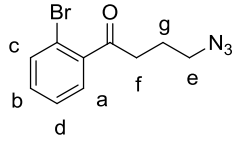
dec.). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 5.71 (s, 2H), 5.86 (s, 2H), 7.64 (s, 1H, H_a). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 59.3, 89.3, 148.4, 155.2, 157.0. LRMS (ESI^+): *m/z* = 338 (100%, $[M+Na]^+$). HRESI-MS: *m/z* = 313.8786 (calcd. for $C_5H_6N_3BrI$, 313.8784).

2-Bromobenzo[c]1,8-naphthyridin-3-amine (41)



To a solution of **43** (12.56 g, 40 mmol) in water (62.5 ml) and 1,4-dioxan (187.5 ml) was added K_2CO_3 (33.6 g), 2-formylphenylboronic acid (9.00 g, 1.5 eq) and $Pd(PPh_3)_4$ (440 mg). The reaction was stirred under nitrogen at 80 °C for 16 hours. The cooled reaction mixture was poured into cold water (750 ml) and filtered. The filtrate was stirred for 1 hour in methanol, then filtered, washed with methanol, and dried to give **41** as a yellow powder (9.58 g, 88%) with ca. 10% of the disubstituted product, which can be easily separated at a later stage. A sample was purified for analysis by recrystallisation from acetonitrile (~20% yield), mp 264 - 267 °C (MeCN). ^{1}H NMR (500 MHz, $(CD_3)_2SO$, 298 K): δ = 6.93 (s, 2H, H_a), 7.66 (dd, ^{3}J = 7.2 Hz, 7.7 Hz, 1H, H_b), 7.87 (dd, ^{3}J = 7.0 Hz, 8.2 Hz, 1H, H_c), 8.15 (d, ^{3}J = 7.7 Hz, 1H, H_d), 8.66 (d, ^{3}J = 8.2 Hz, 1H, H_e), 9.21 (s, 1H, H_f), 9.39 (s, 1H, H_g). ^{13}C NMR (125 MHz, $(CD_3)_2SO$, 298 K): δ = 106.3, 112.0, 122.0, 124.8, 126.7, 129.2, 132.0, 132.6, 136.3, 153.2, 156.8, 157.0. LRMS (ESI^+): *m/z* = 296 (100%, $[M+Na]^+$). HRESI-MS: *m/z* = 273.9976 (calcd. for $C_{12}H_9N_3Br$, 273.9974).

4-Azido-1-(2-bromophenyl)butan-1-one (51)

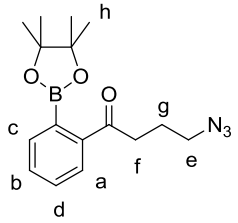


To a freshly-prepared solution of LDA (160 mmol) in THF (250 ml) at -78 °C under nitrogen was added γ-butyrolactone (5.37 ml, 0.44 equivalents). The reaction was stirred for 3 hours at -78 °C, then 2-bromobenzoyl chloride (10.24 ml, 0.49 equivalents) was added dropwise. The reaction was stirred for a further 15 minutes and then quenched with HCl (2 M, 25 ml). The reaction mixture was warmed to room temperature and the THF removed under reduced pressure. The residue was taken up in DCM (250 ml) and washed with HCl (2 M, 2 x 250 ml) and brine (250 ml), then dried over magnesium sulphate and concentrated to give a dark oil (21.44 g).

The oil was suspended in 48% HBr (90 ml) and heated to 80 °C for 1 hour, then cooled and partitioned between water (250 ml) and DCM (250 ml). The layers were separated and the aqueous layer extracted with DCM (100 ml). The combined organic layers were washed with saturated aqueous sodium carbonate (2 x 250 ml), then dried over magnesium sulphate and concentrated to give a dark oil (24.6 g). Traces of DCM were removed under high vacuum (1 torr) to avoid subsequent generation of diazidomethane.

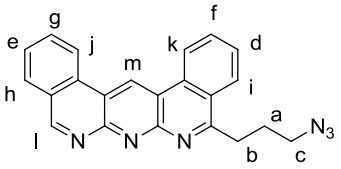
The oil was dissolved in DMF (125 ml) and to this was added NaN_3 (10.37 g, >2 equivalents). The reaction mixture was heated to 80 °C for 1 hour behind a blast shield, then cooled and diluted with EtOAc (250 ml). The organic layer was washed with 5% LiCl (250 ml, 4 x 100 ml) then brine (250 ml), dried over magnesium sulphate and concentrated *in vacuo* to a dark oil (20.28 g). The oil was purified by dry column vacuum chromatography (95:5 heptane:EtOAc) to give **51** as a pale brown oil (13.38 g, 75% over 3 steps). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 2.06 (p, ^3J = 7.0 Hz, 2H, H_g), 3.05 (t, ^3J = 7.0 Hz, 2H, H_f), 3.45 (t, ^3J = 7.0 Hz, 2H, H_e), 7.33 (m, 1H, H_c), 7.43-7.39 (m, 2H, H_b + H_d), 7.64 (m, 1H, H_a). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 23.3, 39.4, 50.7, 118.6, 127.5, 128.3, 131.7, 133.8, 141.5, 203.1. LRMS (ESI^+): *m/z* = 292 ($[M+Na]^+$). HRESI-MS: *m/z* = 268.0083 (calcd. for $C_{10}H_{11}N_3OBr$, 268.0080).

4-Azido-1-[2-(tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl]butan-1-one (52)



To a degassed solution of **51** (13.38 g, 49.9 mmol) in DMSO (250 ml) was added KOAc (14.75 g, 150.3 mmol), B_2Pin_2 (19.25 g, 75.8 mmol), and $Pd(dppf)Cl_2.CH_2Cl_2$ (1.32 g, 1.61 mmol). The reaction mixture was heated to 80 °C for 1 hour, then cooled and extracted with diethyl ether (2 x 250 ml). The combined organic layers were washed with 5% LiCl (4 x 250 ml) and then dried over magnesium sulphate and concentrated *in vacuo*. The residue was purified by dry column vacuum chromatography (90:10 heptane:EtOAc) to give **52** as an orange waxy solid (15.21 g, 97%) mp 50 - 72 °C. ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 1.45 (s, 12H, H_h), 2.07 (qn, ^3J = 6.9 Hz, 2H, H_g), 3.09 (t, ^3J = 7.0 Hz, 2H, H_f), 3.43 (t, ^3J = 6.7 Hz, 2H, H_e), 7.45 (m, 1H, H_c), 7.55 (m, 2H, H_b + H_d), 7.82 (m, 1H, H_a). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 23.4, 24.9, 34.6, 50.8, 83.5, 83.8, 127.5, 129.0, 132.3, 132.5, 140.7, 200.8. LRMS (ESI^+): *m/z* 338 (100%, $[M+Na]^+$). HRESI-MS: *m/z* = 316.1829 (calcd. for $C_{16}H_{32}BN_3O$, 316.1827).

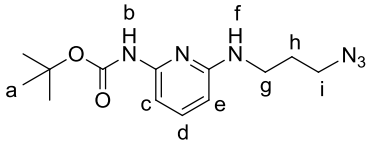
10-(3-Azidopropyl)-1,13,14-Triaza-dibenz[*a,j*]anthracene (29)



To a solution of **41** (2.192 g, 8 mmol) in water (20 ml) and 1,4-dioxan (60 ml) was added **52** (7.50 g, 24 mmol), K_2CO_3 (9.66 g, 69.9 mmol), and $Pd(dppf)Cl_2.CH_2Cl_2$ (200 mg, 0.244 mmol). The reaction was heated to 80 °C for 16 hours, then the hot mixture was poured into a layer of water (500 ml) atop a pad of celite in a sinter funnel (the precipitate is a tacky solid and must not be allowed to settle in a flask). Suction was applied and the celite washed with further water (250 ml) then pulled dry under vacuum for 15 minutes. The celite pad was then triturated with 50:50

MeOH:DCM (500 ml) and filtered. The celite was washed with 50:50 MeOH:DCM (2 x 250 ml) and the combined filtrates were concentrated *in vacuo* to give a dark solid (4.56 g). The solid was purified by dry column vacuum chromatography (THF/heptane, 5 x 100 ml, then THF, 15 x 100 ml) to give **29** as a brown solid (843 mg, 29%) mp 112 - 119 °C (THF). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 2.49 (qn, ^3J = 7.2 Hz, 2H, H_a), 3.60 (t, ^3J = 7.4 Hz, 2H, H_b), 3.67 (t, ^3J = 6.8 Hz, 2H, H_c), 7.85 (m, 2H, H_d + H_e), 8.02 (m, 2H, H_f + H_g), 8.16 (d, ^3J = 7.8 Hz, 1H, H_h), 8.38 (d, ^3J = 8.1 Hz, 1H, H_i), 8.85 (m, 2H, H_j + H_k), 9.63 (s, 1H, H_l), 10.05 (s, 1H, H_m). ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 26.8, 32.6, 51.4, 117.9, 117.9, 122.2, 122.9, 125.4, 126.2, 126.5, 126.9, 128.5, 128.7, 129.4, 131.4, 131.8, 132.4, 132.4, 153.0, 153.9, 159.3, 167.0; LRMS (ESI^+): *m/z* = 365 (100%, M+1). HRESI-MS: *m/z* = 365.1512 (calcd. for $C_{22}H_{17}N_6$, 365.1509).

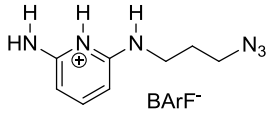
Tert-butyl *N*-(6-[(3-azidopropyl)amino]pyridin-2-yl)carbamate (39)



To a solution of 3-azidopropanol[19] (1.30 g, 12.9 mmol) in dry CH_2Cl_2 (50 mL) was added Dess-Martin periodinane (5.47 g, 13.6 mmol) and the resultant suspension was stirred for 1 h at room temperature under an atmosphere of N_2. A solution of *tert*-butyl (6-aminopyridin-2-yl)carbamate[25] (2.70 g, 12.9 mmol) in CH_2Cl_2 (10 mL) was then added and the resultant mixture was stirred for 10 minutes. $NaBH(OAc)_3$ (5.45 g, 24.8 mmol) was then added in portions and the resultant suspension was stirred for 1 h at room temperature. The crude mixture was diluted with CH_2Cl_2 (100 mL) and carefully added to saturated aqueous $NaHCO_3$ (200 mL). The organic phase was separated and the aqueous layer was washed with further CH_2Cl_2 (2 x 50 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure to furnish a dark brown solid. The crude material was purified using flash column chromatography (1% EtOAc in CH_2Cl_2) to give **39** as a colourless oil (2.56 g, 68%). R*f* = 0.29 (1% EtOAc in CH_2Cl_2); ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 1.52 (s, 9H, H_a), 1.88 (tt, $J^1 = J^2$ = 6.6 Hz, 2H, H_h), 3.37 (m, 2H, H_g), 3.41 (t, ^3J = 6.0 Hz, 2H,

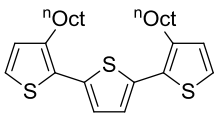
H_i), 4.39 (m, 1H, H_f), 6.07 (d, J = 8.0 Hz, 1H, H_e), 6.94 (s, 1H, H_b), 7.16 (d, ^3J = 8.0 Hz, 1H, H_c), 7.40 (t, ^3J = 8.0 Hz, 1H, H_d); ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 28.3, 28.9, 39.1, 49.2, 80.6, 100.3, 101.4, 139.5, 150.4, 152.2, 157.4; LRMS (ESI^+): *m/z* = 293 (100%, $[M+H]^+$). HRESI-MS: *m/z* = 293.1722 (calcd. for $C_{13}H_{21}O_2N_6$, 293.1721).

2-amino-6-[(3-azidopropyl)amino]pyridin-1-ium tetrakis[(3.5-trifluoromethyl)phenyl]borate (28)



To a solution of **39** (210 mg, 0.718 mmol) in CH_2Cl_2 (4 mL) was added trifluoroacetic acid (4 mL), and the resultant solution was stirred for 16 hours. The reaction mixture was concentrated, taken up in CH_2Cl_2 (10 mL), then washed with water (10 mL) and saturated aqueous $NaHCO_3$ (10 mL). The organics were dried over $MgSO_4$ and concentrated. The solid was dissolved in a solution of HCl in MeOH (1 M, 5 mL) and to this was added NaBArF (620 mg, 0.718 mmol, 1 eq.). The solution was evaporated to dryness, then taken up in CH_2Cl_2, filtered, and evaporated to give **28** as an off-white oil (394 mg, 52%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 1.65 (m, 2H), 3.12 (m, 2H), 3.31 (m, 2H), 5.25 (br.s, 1H), 5.70 (m, 2H), 7.11 (s, 1H), 7.29 (m, 1H), 7.37 (s, 4H), 7.54 (s, 8H), 7.58 (br.s, 1H), 10.5 (br.s, 1H); ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 27.4, 40.7, 49.4, 97.1, 117.6, 123.4, 125.6, 127.7, 128.8, 129.1, 134.7, 146.9. LRMS (ESI^+): *m/z* = 193 [100%, M^+]. LRMS (ESI^-): m/z = 863 [100%, M^-].

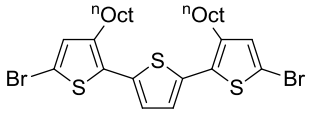
2,5-Bis[2-(*n*octyl)-1-thienyl]thiophene (61)



$Pd(PPh_3)_4$ (0.9 mmol, 1.04 g) was added to a solution of 2-bromo-3-*n*octylthiophene (19.0 mmol, 5.23 g) and 2,5- bis(trimethylstannyl)thiophene (9.3 mmol, 3.80 g) in dimethylformamide (80 ml). The resulting yellow suspension was heated to 100 °C for 16 h. After cooling, the reaction mixture was filtered through alumina and

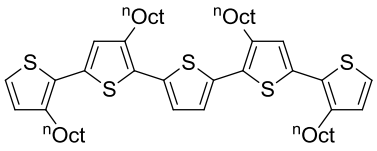
concentrated under reduced pressure. The resulting crude mixture was purified by flash column chromatography (40/60 petroleum spirit) to give **61** as a yellow oil (3.84 g, 87%). ^{1}H NMR (400 MHz, $CDCl_3$, 298 K): δ = 0.87 (m, 6H), 1.19 – 1.43 (m, 20H), 1.65 (m, 4H), 2.78 (t, *J* = 7.9 Hz, 4H), 6.94 (d, *J* = 5.2 Hz, 2H), 7.05 (s, 2H), 7.17 (d, *J* = 5.20 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$, 298 K): δ = 14.1, 22.7, 29.3, 29.5, 29.6, 30.8, 31.9, 123.7, 126.1, 130.1, 130.4, 136.1, 139.7.

2,5-Bis[2-(*n*octyl)-5-(bromo)-1-thienyl]thiophene (62)



61 (8.0 mmol, 3.76 g) and chloroform (10 mL) were dissolved in acetic acid (18 ml) and the resulting solution was cooled to 0 °C. A solution of bromine in chloroform (46.6 mL, 0.35 M) was added drop-wise over 75 minutes. The reaction mixture was quenched with aqueous sodium hydroxide solution (50 mL, 2 M). The mixture was extracted with chloroform (3 x 50 mL) and the combined organic fractions were dried (sodium sulfate) and concentrated under reduced pressure. The resulting crude oil was purified using flash column chromatography (40/60 petroleum spirit) to give **62** as a yellow oil (3.67 g, 73%). ^{1}H NMR (400 MHz, $CDCl_3$, 298 K): δ = 0.89 (t, *J* = 6.80 Hz, 6H), 1.19 – 1.41 (m, 20 H), 1.61 (m, 4H), 2.70 (t, *J* = 7.86 Hz, 4H), 6.90 (s, 2H), 6.98 (s, 2H); ^{13}C NMR (100 MHz, $CDCl_3$, 298 K): δ = 14.6, 22.7, 29.3, 29.5, 29.5, 30.6, 31.7, 110.6, 125.8, 132.0, 132.1, 135.2, 140.5.

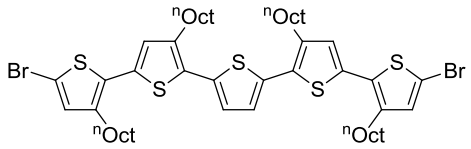
3, 4', 3''', 3''''-Tetraoctyl-2, 5':2', 2'':5'', 2''': 5''', 2''''-quinquethiophene (63)



To a solution of 2-Bromo-3-*n*octylthiophene (17.5 mmol, 4.81 g) in tetrahydrofuran (20 mL) cooled to -8 °C, a solution of isopropyl magnesium chloride in tetrahydrofuran (9 mL, 2 M) was added drop-wise. The resulting solution was allowed

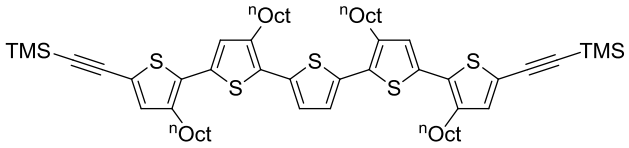
to warm to r.t. and held at temperature for 3 h. The resulting solution was added drop-wise over 25 minutes to a second reaction vessel containing **62** (4.16 mmol, 2.62 g) and 1,3-Bis[(diphenylphosphino)propane] dichloronickel(II) (0.58 mmol, 316 mg) in tetrahydrofuran (15 mL) at 40 °C. The resulting reaction mixture was held at temperature for 16 h after which it was quenched with aqueous acetic acid (25 mL, 20% v/v) and extracted with dichloromethane (50 mL x 2). The combined organics were washed with brine, dried (sodium sulphate) and concentrated under reduced pressure. The resulting crude red oil was purified using flash column chromatography (40/60 petroleum spirit) to give **63** as a yellow oil (2.45 g, 68%). ^{1}H NMR (400 MHz, $CDCl_3$, 298 K): δ = 0.88 (m, 12H), 1.20 – 1.46 (m, 40H), 1.68 (m, 8H), 2.79 (m, 8H), 6.93 (d, *J* = 5.3 Hz, 2H), 6.95 (s, 2H), 7.09 (s, 2H), 7.17 (d, *J* = 5.23 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$, 298 K): δ = 14.1, 22.7, 29.3, 29.5, 29.5, 29.6, 29.6, 29.6, 30.7, 30.8, 32.0, 123.5, 125.4, 129.4, 130.1, 130.2, 130.4, 134.3, 135.8, 139.6, 139.9.

5,5’’’’-Dibromo-3, 4’, 3’’’, 3’’’’-tetraoctyl-2, 5’:2’, 2’’:5’’, 2’’’: 5’’’, 2’’’’-quinquethiophene (64)



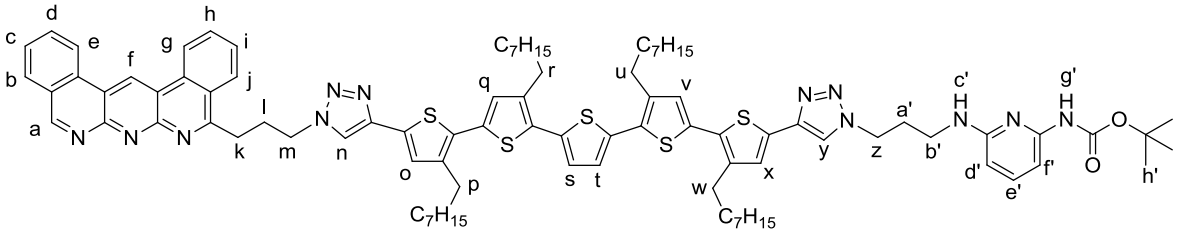
63 (2.76 mmol, 2.38 g) and acetic acid (40 mL) were dissolved in chloroform (12 mL) and cooled to 0 °C. A solution of bromine in chloroform (10 mL, 0.57 M) was added drop-wise over 65 minutes. The reaction mixture allowed to warm to r.t. over 3.5 h before it was quenched with an aqueous sodium hydroxide solution (50 mL, 2M). The mixture was extracted with dichloromethane (50 mL x 2) and the combined organic fractions were dried (sodium sulfate) and concentrated under reduced pressure. The resulting crude oil was purified using flash column chromatography (40/60 petroleum spirit) give **64** as a yellow oil (2.59 g, 92%). ^{1}H NMR (400 MHz, $CDCl_3$, 298 K): δ = 0.94 (m, 12H), 1.26 – 1.51 (m, 40H), 1.61 – 1.78 (m, 8H), 2.76 (t, *J* = 7.70, 4H), 2.82 (t, *J* = 7.70, 4H), 6.91 (s, 2H), 6.93 (s, 2H), 7.11 (s, 2H); ^{13}C NMR (100 MHz, $CDCl_3$, 298 K): δ = 14.2, 22.8, 29.3, 29.4, 29.4, 29.5, 29.6, 29.6, 29.7, 30.7, 30.7, 30.9, 32.0, 110.4, 126.0, 129.1, 130.8, 132.1, 132.8, 133.0, 135.7, 139.9, 140.2.

5,5’’’’-*Bis*[(trimethylsilyl)ethynyl]-3,4’,3’’’,3’’’’-tetraoctyl-2,5’:2’,2’’:5’’,2’’’:5’’’,2’’’’-quinquethiophene (65)



Tetrakis(triphenylphosphine)pal-ladium(0) (59 μmol, 68 mg) and copper(I) iodide (0.19 mmol, 23 mg) were added to a solution of **64** (0.59 mmol, 0.6 g) and trimethylsilylethylene (1.47 mmol, 0.14 g) in toluene (6 mL) and triethylamine (6 mL). The reaction mixture was heated to 80 °C and held at temperature for 16 h. After cooling, the crude reaction mixture was filtered through celite and concentrated under reduced pressure. The resulting crude mixture was purified by flash column chromatography (40/60 petroleum spirit) to give **65** as an orange oil (1.43 g, 74%). ^{1}H NMR (400 MHz, $CDCl_3$, 298 K): δ = 0.31 (s, 18H), 0.93 (t, *J* = 6.80 Hz, 12H), 1.26 – 1.52 (m, 40H), 1.63 – 1.78 (m, 8H), 2.77 (t, *J* = 7.90 Hz, 4H), 2.83 (t, *J* = 7.9 Hz, 4H), 7.00 (s, 2H), 7.11 (s, 2H), 7.13 (s, 2H); ^{13}C NMR (100 MHz, $CDCl_3$, 298 K): δ = 0.0, 14.1, 22.4, 29.2, 29.3, 29.3, 29.4, 29.4, 29.4, 29.5, 29.6, 30.4, 30.6, 31.9, 97.6, 99.7, 110.0, 120.6, 126.1, 129.1, 130.7, 132.4, 133.4, 135.7, 139.3, 140.1.

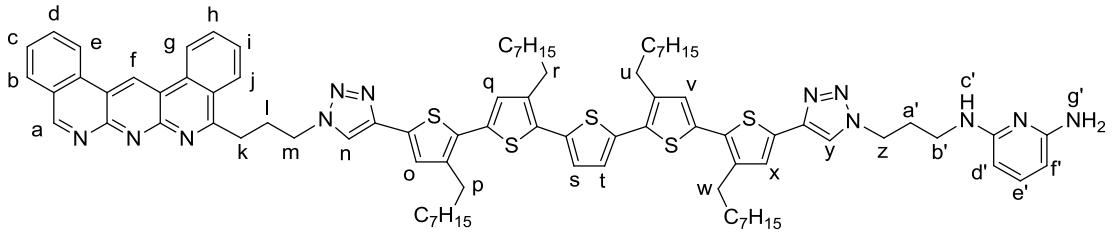
AAA-pentathiophene-DAD-Boc (66)



Potassium hydroxide (15 mg, 0.27 mmol) dissolved in methanol (0.5 mL) was added drop-wise to a solution of **65** (27.5 μmol, 28.9 mg) in toluene (1 mL). The resulting solution was stirred at rt for 1.5 h before water (1.5 mL) was added and the mixture extracted with dichloromethane (5 mL x 2). The organics were washed with water (5 mL x 2) and brine before being dried (sodium sulfate) and concentrated under reduced pressure.

To the crude reaction mixture was added **39** (8.1 mg, 27.5 μmol), **29** (10.0 mg, 27.5 μmol), 20% v/v methanol/dichloromethane (2 mL), triethylamine (*c.a.* 50 mg) and tetrakis(acetonitrile-*N*)copper(I) hexafluorophosphate(V) (2 mg, 5.5 μmol). The resulting mixture was stirred at rt for 3 h before solvents were removed under reduced pressure. The resulting crude mixture was purified by flash column chromatography (0.5 % v/v – 8% v/v 60 methanol/dichloromethane) to give **66** as an orange solid (10.5 mg, 24.4%). ^{1}H NMR (500 MHz, $CDCl_3$, 298 K): δ = 0.88 (m, 12H, alkyl-*H*), 1.21-1.45 (m, 32H, alkyl-*H*), 1.56-1.74 (m, 16 H, alkyl-*H*), 2.26 (m, 2H, $H_{a'}$), 2.59 (t, 2H, J = 6.7 Hz, H_p), 2.80 (m, 6H, H_{r+u+w}), 2.92 (m, 2H, H_l), 3.38 (m, 2H, $H_{b'}$), 3.62 (t, 2H, J = 6.5 Hz, H_k), 4.52 (t, 2H, J = 6.6 Hz, H_z), 4.77 (t, 2H, , J = 6.1 Hz, H_m), 6.09 (d, 1H, J = 8.0 Hz, $H_{d'}$), 6.90 (s, 1H, H_q), 7.01 (s, 1H, H_o), 7.04 (s, 1H, H_v), 7.11 (m, 2H, H_{s+t}), 7.25 (s, 1H, H_x), 7.41 (t, 1H, J = 8.0 Hz, $H_{e'}$), 7.68 (s, 1H, H_y), 7.84 (m, 3H, H_{c+i+n}), 8.00 (m, 2H, H_{h+d}), 8.17 (d, 1H, J = 7.8 Hz, H_b), 8.31 (d, 1H, J = 8.1 Hz, H_j), 8.83 (d, 1H, J = 8.0 Hz, H_e), 8.88 (d, 1H, J = 7.7 Hz, H_g), 9.67 (s, 1H, H_f), 10.05 (s, 1H, H_a); ^{13}C NMR (125 MHz, $CDCl_3$, 298 K): δ = 14.1 (x4), 22.7 (x4), 27.3, 28.3, 29.3, 29.3, 29.3 (x2), 29.4 (x2), 29.4 (x2), 29.4 (x2), 29.5 (x2), 29.5 (x2), 29.5 (x2), 29.6 (x2), 29.6, 29.6, 30.0, 30.3, 30.5, 30.6, 31.9, 32.1, 38.5, 47.9, 50.1, 80.8, 100.3, 102.0, 118.0, 119.2, 120.0, 122.2, 122.9, 125.3, 125.9, 126.0, 126.2, 126.4, 126.8, 127.0, 127.2, 128.5, 128.6, 128.8, 128.8, 129.5, 129.8, 130.1, 130.2, 130.3, 130.4, 130.5, 131.6, 131.9, 132.4, 132.4, 133.9, 134.1, 135.6, 135.9, 138.9, 139.9, 139.9, 140.0, 140.2, 142.3, 142.5, 142.7, 152.1, 152.8, 154.0, 159.4, 161.9, 163.0, 166.3; ESI-MS: *m/z* = 784 (100%, $[M+2H]^{2+}$); HRESI-MS: *m/z* = 783.8904 (calculated for $C_{91}H_{114}N_{12}O_2S_5$, 783.8905).

AAA-pentathiophene-DAD (59)



Neat trifluoroacetic acid (5 mL) was added dropwise to a solution of **66** (62 mg, 39.6 μmol) in CH_2Cl_2 (5 mL) at 0 ºC under an atmosphere of N_2. After 3 hours the solution

was diluted with CH_2Cl_2 (50 mL) and then carefully quenched with saturated aqueous $NaHCO_3$ (50 mL). The organic phase was separated and the aqueous layer was washed with further CH_2Cl_2 (2 x 25 mL). The combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure to furnish a dark orange solid. The crude residue was purified using flash column chromatography (5% MeOH/95% CH_2Cl_2) to give **59** as an orange oil (37 mg, 64%). ^{1}H NMR (500 MHz, CD_2Cl_2/d_3-MeOD (19:1), 298 K): δ = 0.92 (m, 12 H, alkyl-*H*), 1.06-1.54 (m, 40H, alkyl-*H*), 1.61 (m, 2H, alkyl-*H*), 1.74 (m, 6H, alkyl-*H*), 2.27 (m, 2H, $H_{a'}$), 2.61 (t, 2H, *J* = 6.7 Hz, H_p), 2,85 (m, 6H, H_{r+u+w}), 2.90 (m, 2H, H_l), 3.33 (t, 2H, *J* = 6.5 Hz, $H_{b'}$), 3.67 (t, 2H, *J* = 6.5 Hz, H_k), 4.54 (t, 2H, *J* = 6.5 Hz, H_z), 4.76, 2H, *J* = 6.5 Hz, H_m), 5.81 (m, 1H, $H_{d'}$), 5.88 (m, 1H, $H_{f'}$), 6.93 (s, 1H, H_q), 7.04 (s, 1H, H_o), 7.08 (s, 1H, H_v), 7.18 (m, 2H, H_{s+t}), 7.27 (s, 1H, H_x), 7.28 (m, 2H, $H_{e'}$), 7.86 (s, 1H, H_y), 7.89 (m, 2H, H_{c+i}), 8.05 (s, 1H, H_n), 8.06 (m, 2H, H_{h+d}), 8.23 (d, 1H, *J* = 7.4 Hz, H_b), 8.40 (d, 1H, *J* = 7.4 Hz, H_j), 8.90 (d, 1H, *J* = 7.4 Hz, H_e), 8.98 (d, *J* = 7.4 Hz, H_g), 9.57 (s, 1H, H_f), 10.13 (s, 1H, H_a); ^{13}C NMR (125 MHz, CD_2Cl_2/d_3-MeOD (19:1), 298 K): δ = 13.8 (x4), 22.6 (x4), 27.5, 29.3, 29.3 (x2), 29.3 (x2), 29.3 (x2), 29.4 (x2), 29.4 (x2), 29.4 (x2), 29.6 (x2), 29.7 (x2), 29.6 (x2), 29.6, 30.3, 30.5, 30.6, 30.6, 31.9, 32.5, 38.8, 48.0, 50.2, 94.5, 96.7, 118.1, 118.2, 120.1, 120.7, 122.4, 123.1, 125.0, 125.0, 125.9, 126.0, 126.0, 126.4, 126.9, 127.2, 127.8, 128.5, 128.7, 128.9, 129.5, 129.7, 130.0, 130.3, 130.3, 131.8, 132.2, 132.3, 132.3, 133.8, 133.8, 135.6, 135.7, 140.0, 140.1, 140.2, 140.4, 141.0, 141.0, 141.1, 141.9, 142.3, 152.2, 153.0, 156.1, 156.6, 159.4, 167.1; ESI-MS: *m/z* = 784 (100%, $[M+H]^{2+}$); mass identity confirmed by isotopic distribution (Figure 4.15).

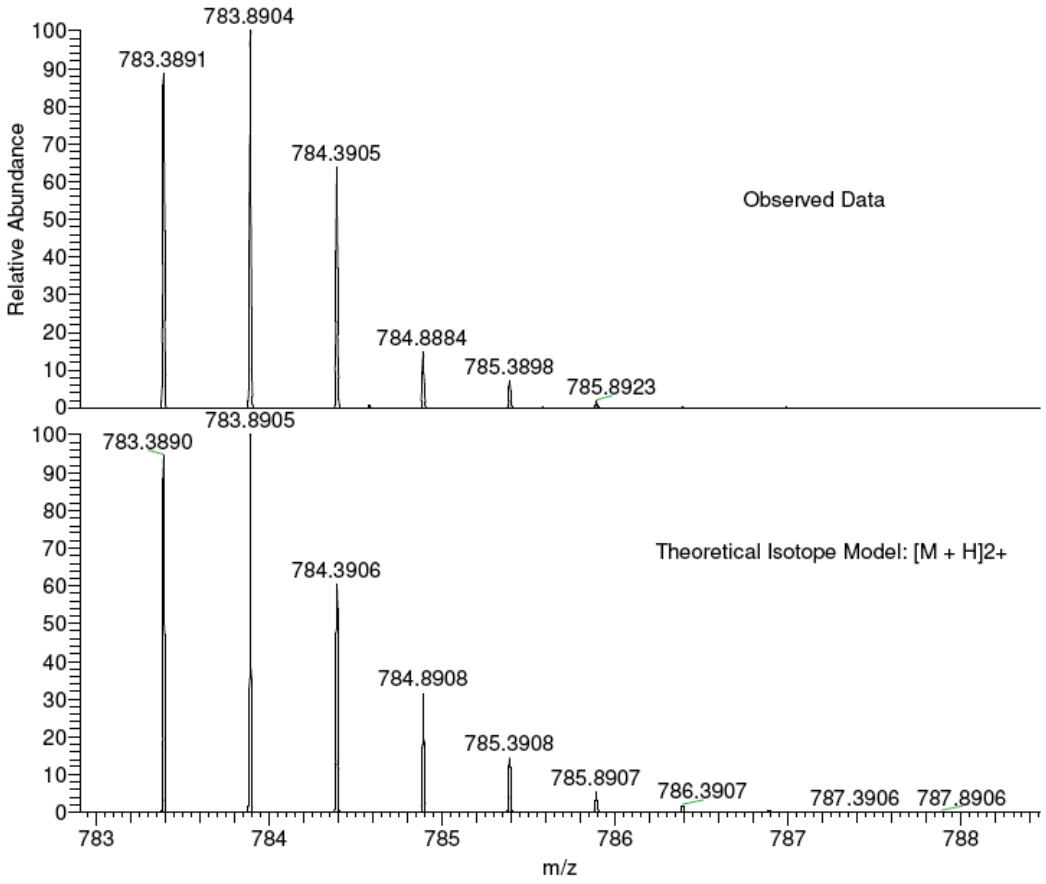


Figure 4.15 | Isotopic distribution patterns for compound 59, $[M+H]^{2+}$. Real (top) and calculated (bottom). Although the accurate masses also match, this is not sufficient for molecular formula confirmation when mass > 1,000 Da.

4.8 References

(1) de Greef, T. F. A.; Meijer, E. W. *Nature* **2008**, *453*, 171.

(2) Brunsveld, L.; Folmer, B. J. B.; Meijer, E. W.; Sijbesma, R. P. *Chem. Rev.* **2001**, *101*, 4071.

(3) De Greef, T. F. A.; Smulders, M. M. J.; Wolffs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109*, 5687.

(4) Wilson, A. J. *Soft Matter* **2007**, *3*, 409.

(5) Kotera, M.; Lehn, J.-M.; Vigneron, J.-P. *J. Chem. Soc., Chem. Commun.* **1994**, 197.

(6) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008.

(7) Sijbesma, R. P.; Beijer, F. H.; Brunsveld, L.; Folmer, B. J. B.; Hirschberg, J. H. K. K.; Lange, R. F. M.; Lowe, J. K. L.; Meijer, E. W. *Science* **1997**, *278*, 1601.

(8) Folmer, B. J. B.; Sijbesma, R. P.; Versteegen, R. M.; van der Rijt, J. A. J.; Meijer, E. W. *Adv. Mater.* **2000**, *12*, 874.

(9) Hentschel, J.; Kushner, A. M.; Ziller, J.; Guan, Z. *Angew. Chem., Int. Ed. Engl.* **2012**, 10561.

(10) Huerta, E.; Cequier, E.; Mendoza, J. d. *Chem. Commun.* **2007**, 5016.

(11) De Greef, T. F. A.; Smulders, M. M. J.; Wolffs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109*, 5687.

(12) J. B. Folmer, B.; Cavini, E. *Chem. Commun.* **1998**, 1847.

(13) Corbin, P. S.; Zimmerman, S. C. *J. Am. Chem. Soc.* **1998**, *120*, 9710.

(14) Park, T.; Zimmerman, S. C.; Nakashima, S. *J. Am. Chem. Soc.* **2005**, *127*, 6520.

(15) Li, Y.; Park, T.; Quansah, J. K.; Zimmerman, S. C. *J. Am. Chem. Soc.* **2011**, *133*, 17118.

(16) Leigh, D. A.; Djurdjevic, S.; McNab, H.; Parsons, S.; Teobaldi, G.; Zerbetto, F. *J. Am. Chem. Soc.* **2007**, *129*, 476.

(17) Leigh, D. A.; Blight, B. A.; Camara-Campos, A.; Djurdjevic, S.; Kaller, M.; McMillan, F. M.; McNab, H.; Slawin, A. M. Z. *J. Am. Chem. Soc.* **2009**, *131*, 14116.

(18) Braxmeier, T.; Demarcus, M.; Fessmann, T.; McAteer, S.; Kilburn, J. D. *Chem.--Eur. J.* **2001**, *7*, 1889.

(19) Pak, J. K.; Hesse, M. *J. Org. Chem.* **1998**, *63*, 8200.

(20) Mishra, A.; Ma, C.-Q.; Bäuerle, P. *Chem. Rev.* **2009**, *109*, 1141.

(21) Colella, N. S.; Zhang, L.; Briseño, A. L. *Material Matters* **2012**, *7*, 18.

(22) Gall, M., University of Edinburgh, 2012.

(23) Pedersen, D. S.; Rosenbohm, C. *Synthesis* **2001**, *2001*, 2431.

(24) Boudakian, M. M.; Olin Corp.: US3849429, 1974.

(25) Cooke, G.; Caldwell, S. T.; Hewage, S. G.; Mabruk, S.; Rabani, G.; Rotello, V.; Smith, B. O.; Subramani, C.; Woisel, P. *Chem. Commun.* **2008**, 4126.

(26) Elhaïk, J; Pask, C. M.; Kilner, C. A.; Halcrow, M. A. *Tetrahedron* **2007**, *63*, 2, 291.

Appendix I: Published papers

This appendix contains the text of published papers deriving from work reported in this thesis, as they appeared in print.

(1) An AAAA-DDDD Quadruple Hydrogen Bond Array, Blight, B. A.; Hunter, C. A.; Leigh, D. A.; McNab, H.; Thomson, P. I. T. *Nature Chemistry* **2011**, *3*, 244-248.

ARTICLES
PUBLISHED ONLINE: 21 FEBRUARY 2011 | DOI: 10.1038/NCHEM.987
nature chemistry

An AAAA–DDDD quadruple hydrogen-bond array

Barry A. Blight[1], Christopher A. Hunter[2], David A. Leigh[1]*, Hamish McNab[1†] and Patrick I. T. Thomson[1]

Secondary electrostatic interactions between adjacent hydrogen bonds can have a significant effect on the stability of a supramolecular complex. In theory, the binding strength should be maximized if all the hydrogen-bond donors (D) are on one component and all the hydrogen-bond acceptors (A) are on the other. Here, we describe a readily accessible AAAA–DDDD quadruple hydrogen-bonding array that exhibits exceptionally strong binding for a small-molecule hydrogen-bonded complex in a range of different solvents ($K_a > 3 \times 10^{12}$ M^{-1} in CH_2Cl_2, 1.5×10^6 M^{-1} in CH_3CN and 3.4×10^5 M^{-1} in 10% v/v DMSO/$CHCl_3$). The association constant in CH_2Cl_2 corresponds to a binding free energy (ΔG) in excess of −71 kJ mol^{-1} (more than 20% of the thermodynamic stability of a carbon–carbon covalent bond), which is remarkable for a supramolecular complex held together by just four intercomponent hydrogen bonds.

Multipoint hydrogen-bonding motifs are the cornerstone of recognition processes in biology and are increasingly featuring in the design of multifunctional materials and supramolecular polymers[1–5]. In 1990, Jorgensen suggested that secondary electrostatic interactions play an important role in the stability of arrays of contiguous hydrogen bonds (Fig. 1)[6]. A corollary of this is that having all the hydrogen-bond donor groups (D) in one partner and all the hydrogen-bond acceptor sites (A) in the other should be the arrangement that produces the strongest binding, because all secondary electrostatic interactions between neighbouring hydrogen-bond pairs are attractive. Although the significance of this effect has been questioned based on the results of quantum-mechanical calculations on DNA base pairs[7], it accounts well for the general experimental trends found for triple hydrogen-bonded complexes (typical complex stability ADA–DAD < ADD–DAA < AAA–DDD)[8–14] and can be used quantitatively in empirical methods for predicting complex stabilities[8,15].

The experimental trends are less clear cut for quadruple hydrogen-bond complexes (see, for example, **9·9** and **10·11**, Fig. 2). However, there are far fewer examples of such arrays, and with a small sample set it is difficult to separate the contribution of the arrangement of the hydrogen-bond pairs from other factors that contribute to complex stability, such as the hydrogen-bond acidity/basicity of the functional groups, additional CH···O/N or multipole interactions, entropy effects, solvation, tautomerism and the strength of homodimerization of each partner. Indeed, of the six possible quadruple hydrogen-bond permutations[16–27], only three have been definitively experimentally realized to date: ADAD–DADA (for example, **8·8**)[16], AADD–DDAA (for example, **9·9**)[17–21] and ADDA–DAAD (for example, **10·11**)[22–24] (Fig. 2). The AAAD–DDDA motif remains experimentally unexplored. Claims for an ADAA–DADD (ref. 25) and a cationic AAAA–DDDD$^+$ (ref. 26) complex have been made, but both have anomalously low stability constants ($K_a = 590$ M^{-1} and 530 M^{-1} in $CDCl_3$, respectively) as a result of intramolecular hydrogen bonding that favours rotamers that do not correspond to the desired hydrogen-bonding array (see Supplementary Information). Here we report on a readily accessible AAAA–DDDD$^+$ complex **12·13** that exhibits exceptional complex stability even in hydrogen-bond-disrupting solvent systems.

Results and discussion

Synthesis of 12 and 13. The quadruple hydrogen-bond donor (**12**) and acceptor (**13**) units were each synthesized in three steps from commercially available starting materials (Fig. 3). The DDDD$^+$ component **12** was assembled from 2-aminobenzimidazole **14** via formation of thiourea **15** (CS_2, pyridine, 130 °C, 18 h, 81%),

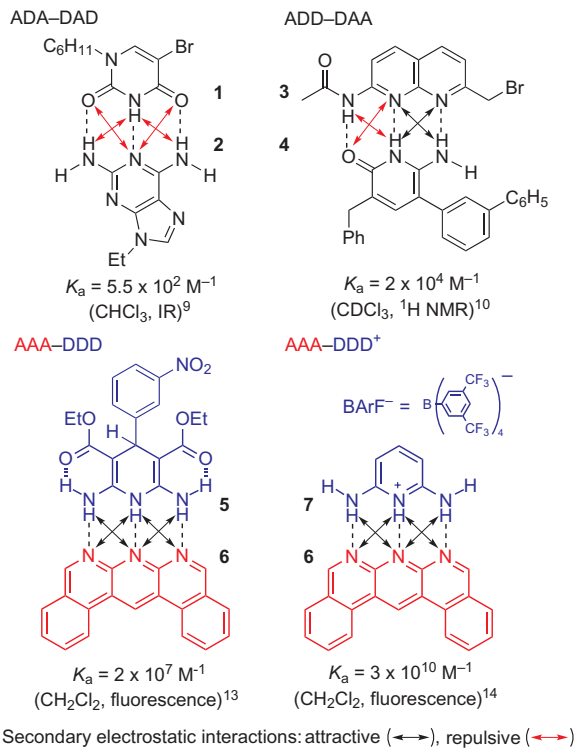


Figure 1 | Examples of triple hydrogen-bond arrays and their association constants in various solvents. Different permutations of contiguous triple hydrogen-bonded complexes ADA–DAD **1·2** (ref. 9), ADD–DAA **3·4** (ref. 10), AAA–DDD **5·6** (ref. 13) and cationic AAA–DDD$^+$ **7·6** (ref. 14). Arrows indicate secondary electrostatic interactions (black, attractive; red, repulsive).

[1]School of Chemistry, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JJ, UK, [2]Department of Chemistry, University of Sheffield, Sheffield S3 7HF, UK; †Deceased. *e-mail: david.leigh@ed.ac.uk

© 2011 Macmillan Publishers Limited. All rights reserved.

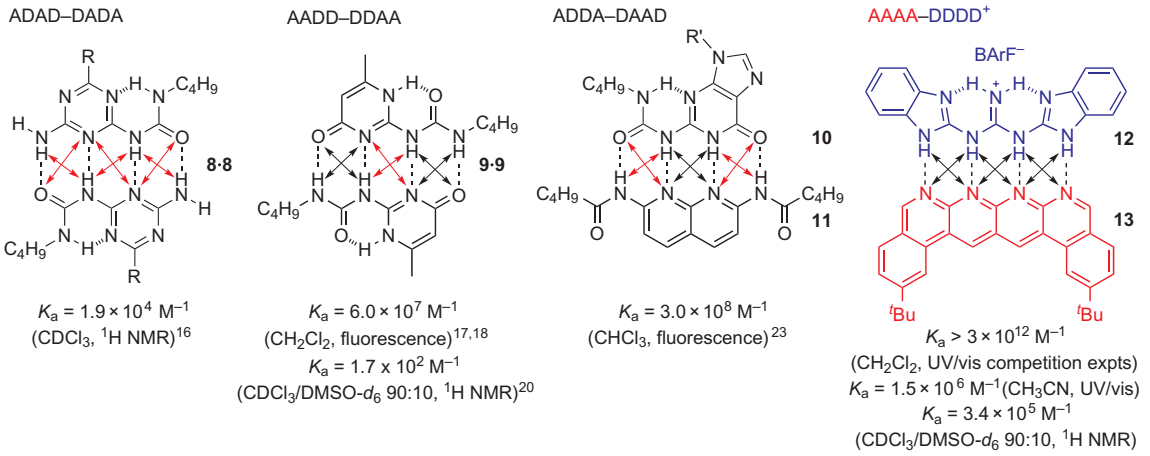


Figure 2 | Examples of quadruple hydrogen-bond arrays and their association constants in various solvents. Different permutations of contiguous quadruple hydrogen-bonded complexes: ADAD–DADA **8·8** (ref. 16), AADD–DDAA **9·9** (refs 17,18,20), ADDA–DAAD **10·11** (ref. 23) and AAAA–DDDD$^+$ **12·13**. Arrows indicate secondary electrostatic interactions (black, attractive; red, repulsive).

conversion to guanidine **16** (HgO, NH_3/MeOH, $CHCl_3$, 25 °C, 3 h, 50%), followed by protonation and ion-exchange (NaBArF, 8 M AcOH, 25 °C, 2 h, 40%) to the weakly coordinating $[B(3,5\text{-}(CF_3)_2C_6H_3)_4]^-$ ($BArF^-$) anion. The AAAA partner, **13**, was synthesized from 2,7-diamino-1,8-naphthyridine **17** via a double bromination with *N*-bromosuccinimide (NBS) to afford **18** (NBS, DMF (dimethylformamide), –10 to 25 °C, 3 h, 60%), followed by a one-pot double Suzuki coupling–cyclization–aromatization procedure[13,14] with boronic acid **19**.

Characterization of complex 12·13. The DDDD$^+$ component **12** has two intramolecular hydrogen bonds to help stabilize tautomers that present three or four hydrogen-bond donors along one edge of the molecule (**12** and **12′**, Fig. 3). The cationic charge should increase the donor strength of the hydrogen-bonding groups compared to neutral systems[12,14]. The AAAA unit **13** is a hexacene system intended to improve its chemical stability compared to underivatized linear arrays of pyridine rings linked through their 2,3/4,5 edges[11,12]. UV/vis titration of **13** with **12** in CH_2Cl_2 ($\sim 5 \times 10^{-5}$ M) showed a decrease in the absorption spectrum of **13** ($\lambda_{max} = 426$ nm) accompanied by the appearance of a new species bathochromically shifted by 11 nm (Fig. 4). The complex was characterized by electrospray ionization mass spectrometry (ESI-MS) ($[\mathbf{12{\cdot}13}]^+$ $m/z = 736.44$; Supplementary Fig. S7) and ^{1}H NMR (nuclear magnetic resonance) spectroscopy (Fig. 5). Particularly noteworthy are the large downfield shifts (up to 10 ppm) of the NH protons of **12** upon complex formation with **13**, and the upfield shifts of the benzimidazole CH protons as the NH bonds become more polarized through hydrogen bonding. The broad NH signals in free **12** may be a consequence of the interconversion of the tautomers shown in Fig. 3 (ref. 28). In contrast, well-resolved signals for three different types of NH protons (H_{abc}) are observed in the spectrum of **12·13**, as expected for the DDDD$^+$ tautomer of **12**. Rotating frame Overhauser effect spectroscopy (ROESY) experiments showed through-space interactions between the H_C protons of **13** and the H_d protons of **12** (Supplementary Fig. S4).

Spectroscopic evaluation of the stability of AAAA–DDDD$^+$ complex 12·13 in dichloromethane. The binding constant of **12·13** proved too strong to be measured directly by UV/vis or ^{1}H NMR titrations (these methods are generally only useful for binding constants up to $\sim 1 \times 10^5$ M^{-1})[29], and **13** proved photochemically unstable at the high light intensities required to

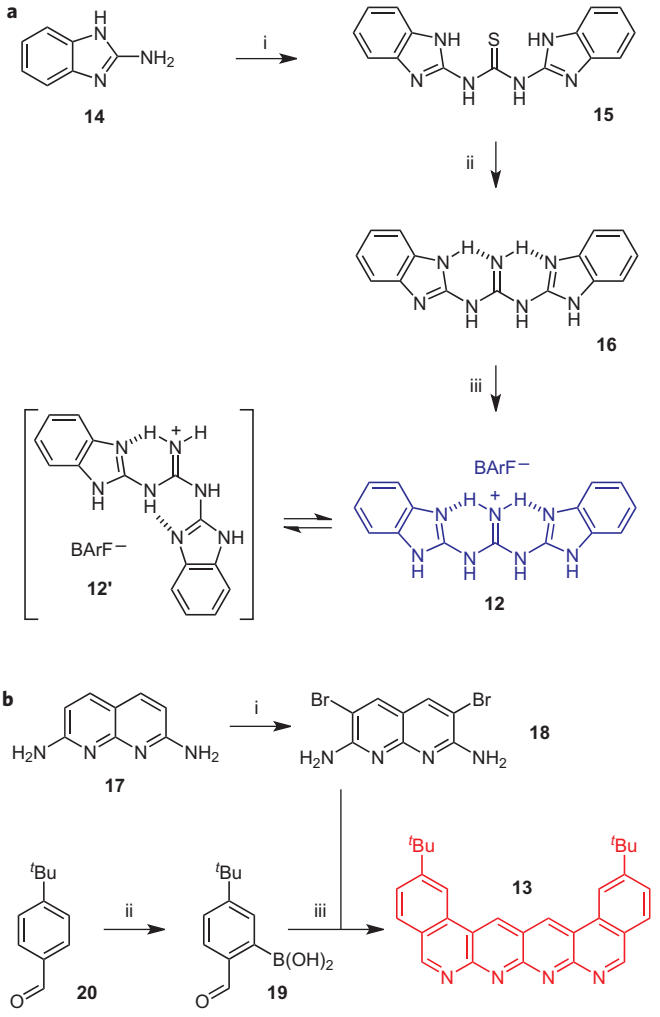


Figure 3 | Synthetic routes to DDDD$^+$ 12 and AAAA 13. a,**b**, Synthesis of DDDD$^+$ **12** ($BArF^-$ salt) (**a**) and AAAA **13** (**b**). Reagents and conditions for **a**: (i) CS_2, pyridine, 130 °C, 18 h, 81%; (ii) HgO, NH_3/MeOH, $CHCl_3$ 25 °C, 3 h, 50%; (iii) NaBArF, 8 M AcOH, 25 °C, 2 h, 40%. Reagents and conditions for **b**: (i) NBS, DMF, –10 to 25 °C, 3 h, 60%; (ii) $NHCH_3(CH_2CH_2)N(CH_3)_2$, BuLi then $B(OMe)_3$, THF, –78 to 25 °C, 16 h, 36%; (iii) Na_2CO_3, H_2O/DME (dimethoxyethane) 1:1, 80 °C, 1.5 h, 27%.

© 2011 Macmillan Publishers Limited. All rights reserved.

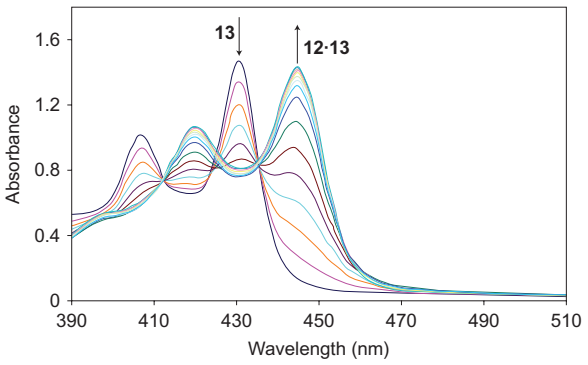


Figure 4 | UV/Vis titration of AAAA 13 with DDDD$^+$ 12 in CH_2Cl_2. UV/vis spectra of **13** ($\sim 5 \times 10^{-5}$ M) on addition of **12** (0 → 5 equiv.), maintaining the concentration of **13** constant, in CH_2Cl_2 at 298 K. Changes in absorbance reflect changes in the amount of **13** and **12·13** present during the titration experiment and differences in their UV/vis absorption spectra.

measure changes in fluorescence at low concentrations ($\sim 1 \times 10^{-10}$ M). Therefore competition experiments were performed in CH_2Cl_2 to compare the association constant for **12·13** to that of known complex AAA–DDD$^+$ **7·6**. Complex **7·6** was titrated with **12** (followed by changes in the UV/vis spectrum) to give $K_a = 1 \times 10^{10}$ M^{-1} ($\Delta G = -57.1$ kJ mol^{-1}) for AAA–DDDD$^+$ complex **12·6**, slightly less than that of **7·6** ($K_a = 3 \times 10^{10}$ M^{-1}), probably as a result of the guanidinium DDDD$^+$ unit being less hydrogen-bond acidic than the pyridinium DDD$^+$ species. Density function calculations (B3LYP, 6-31G*, see Supplementary Information) carried out on **7** and **12** revealed that the pyridinium NH of **7** carries the largest positive partial charge (+667 kJ mol^{-1}), whereas its amino-protons carry an electrostatic potential of +605 kJ mol^{-1}. In **12**, the guanidinium protons have an electrostatic potential of +660 kJ mol^{-1} and the benzimidazole NH protons +565 kJ mol^{-1}. (Note also that **12** may bind to the AAA system to a significant extent as the alternative DDD$^+$ tautomer **12′**, shown in black in Fig. 3.)

A UV/vis titration of **12·6** with **13** gave a lower limit for the K_a of **12·13** of 1×10^{11} M^{-1} (the experiment being complicated by 1:1 and 2:1 complexes of **12·13** and the overlapping absorption spectra of **6** and **13**). The reverse UV/vis experiment was also performed (titrating AAAA–DDDD$^+$ **12·13** with a large excess of AAA **6** to liberate **13**) to determine the stability of **12·13** at high concentrations of **6** (Fig. 6). Free **13** began appearing above a threshold of 350 equiv. of **6**, indicating a K_a of $>3 \times 10^{12}$ M^{-1} ($\Delta G < -71$ kJ mol^{-1}) for AAAA–DDDD$^+$ **12·13** in CH_2Cl_2.

Complex stability of AAAA–DDD$^+$ complex 12·13 in other solvent systems. Solvent competition for hydrogen-bonding sites can dramatically influence complex stability. If the thermodynamic contributions of individual hydrogen bonds in a multiply hydrogen-bonded complex are assumed to be additive in free energy, the magnitude of the expected solvent effects can be estimated using Hunter's hydrogen bond parameters[30] in conjunction with the following equation:

$$\Delta G^0 = -\sum_i (\alpha_i - \alpha_S)(\beta_i - \beta_S) + 6\ \text{kJ mol}^{-1} \qquad (1)$$

where α_S and β_S are the hydrogen-bond parameters of the solvent, α_i and β_i are the hydrogen-bond parameters for the solute interaction sites, the sum is over all solute–solute interactions, and 6 kJ mol^{-1} is a constant representing the cost of forming a bimolecular complex in solution.

If we assume that all the hydrogen-bond donor sites on **12** and **7** are similar and comparable to other ammonium and guanidinium cation hydrogen-bond donors that have been characterized experimentally (C.A. Hunter *et al.*, unpublished results), it is possible to estimate stability constants for the **7·6** and **12·13** complexes in dichloromethane and in acetonitrile from equation (1) (see Supplementary Information). The calculated values in dichloromethane, 4×10^9 M^{-1} and 2×10^{13} M^{-1} for three and four hydrogen bonds, respectively, are consistent with the experimental measurements. Changing the solvent to acetonitrile, a solvent in which the strength of solute–solute interactions are significantly reduced[30], lowers the stability constants estimated using equation (1) to 2×10^3 M^{-1} and 4×10^4 M^{-1}, respectively.

UV/vis titration of AAA–DDD$^+$ complex **7·6** gave $K_a = 5.4 \times 10^3$ M^{-1} ($\Delta G = -21.3$ kJ mol^{-1}) in CH_3CN, far less than the K_a of 3.0×10^{10} M^{-1} in CH_2Cl_2 and in good agreement with the value estimated from equation (1). During the titration of **13** with **12** (Fig. 7a), in addition to the 1:1 complex a 2:1 binding mode was also observed (confirmed by a Job plot; see Supplementary Information). The complex stabilities of **12·13** and **13·(12·13)** were determined to be $K_{\mathbf{12\cdot13}} = 1.5 \times 10^6$ M^{-1} ($\Delta G = -35.2$ kJ mol^{-1}) (280× stronger than **7·6** and 50× stronger than the value estimated from equation (1)) and, for the second binding event, $K_{\mathbf{13\cdot(12\cdot13)}} = 2.4 \times 10^5$ M^{-1} ($\Delta G = -30.6$ kJ mol^{-1}). The formation of **13·(12·13)** is probably due to favourable electrostatic interactions between **13** and **12·13**. A ^{1}H NMR dilution experiment of **13** revealed a small π-stacking component ($K_{dim} < 100$ M^{-1}; see Supplementary Information), suggesting that there may be some reinforcing effect that biases the formation of **13·(12·13)** until equimolar

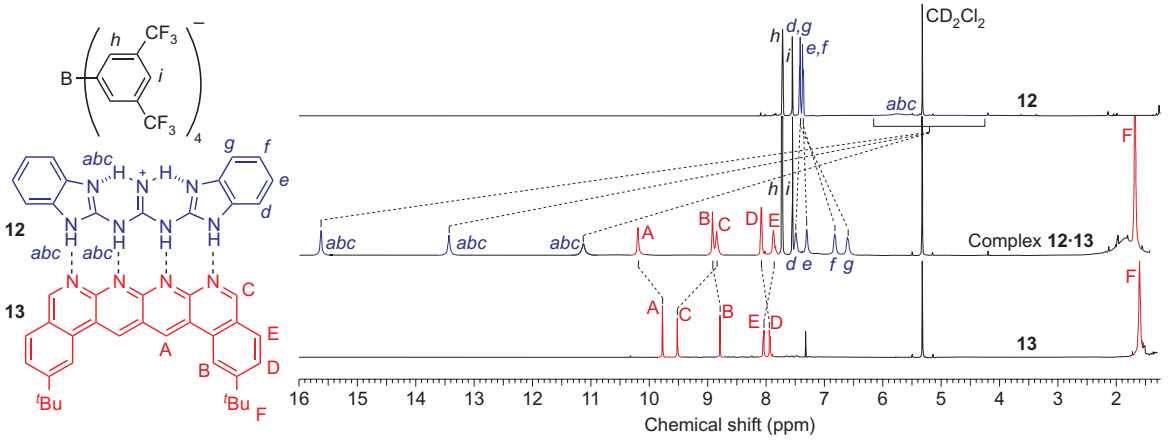


Figure 5 | ^{1}H NMR spectra (500 MHz, CD_2Cl_2, 298 K) of 12 (top), complex 12·13 (middle) and 13 (bottom). Dashed lines show the changes in chemical shift of the resonances in **12** and **13** on formation of complex **12·13**.

© 2011 Macmillan Publishers Limited. All rights reserved.

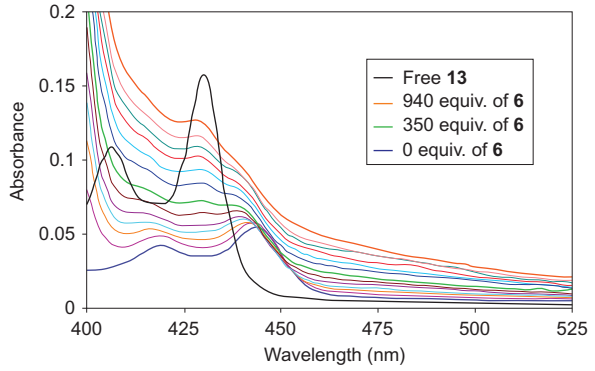


Figure 6 | UV/vis competition experiment in which 13 is displaced from 12·13 by a large excess of 6. UV/vis spectra of **12·13** ($\sim 2 \times 10^{-7}$ M) following addition of **6** (0 → 950 equiv., 50 equiv. aliquots) in CH_2Cl_2 at 298 K. Spectra at 0 equiv. (blue), 350 equiv. (green) and 940 equiv. (orange) of added **6** are shown with thicker linewidths for clarity. For comparison, the spectrum of free **13** (black) is also shown.

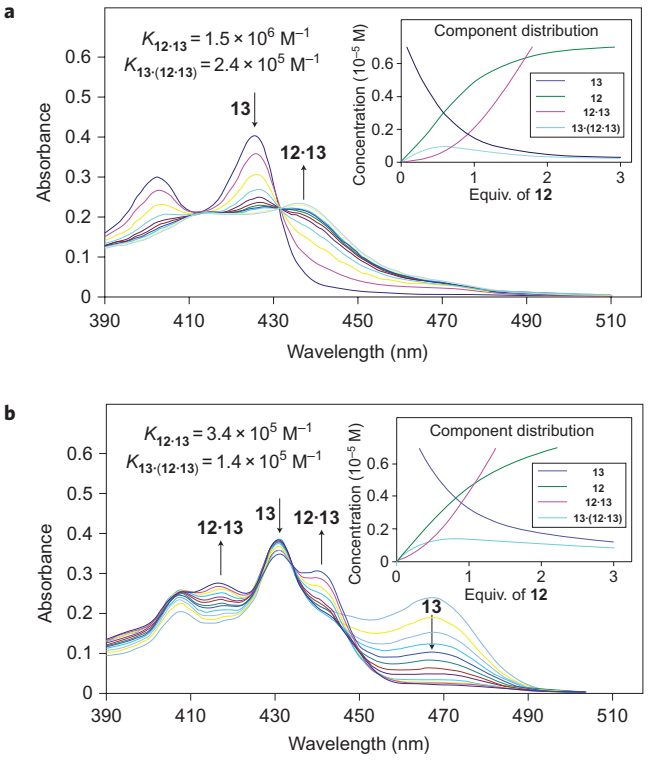


Figure 7 | UV/vis titration experiments in which $K_{12·13}$ and $K_{13·(12·13)}$ are measured by addition of 12 to 13 in CH_3CN or 10% DMSO in $CHCl_3$. **a**, UV/vis spectra of **13** ($\sim 8 \times 10^{-6}$ M) following addition of **12** (0 → 5 equiv.), while maintaining the concentration of **13** constant, in CH_3CN at 298 K. Component distribution over the course of the titration experiment is also illustrated (inset). Association constants for **12·13**, modelled with a 2:1 equilibrium, are $K_{12·13} = 1.5 \times 10^6$ M^{-1} ($\Delta G = -35.2$ kJ mol^{-1}) for **12·13** and $K_{13·(12·13)} = 2.4 \times 10^5$ M^{-1} ($\Delta G = -30.6$ kJ mol^{-1}) for the binding of a second AAAA unit to this complex to form **13·(12·13)**. **b**, UV/vis spectra of **13** ($\sim 1 \times 10^{-5}$ M) following addition of **12** (0 → 4.5 equiv.), while maintaining the concentration of **13** constant, in a solution of 10% DMSO in $CHCl_3$ at 298 K. Association constants are $K_{12·13} = 3.4 \times 10^5$ M^{-1} ($\Delta G = -31.6$ kJ mol^{-1}) and $K_{13·(12·13)} = 1.4 \times 10^5$ M^{-1} ($\Delta G = -29.4$ kJ mol^{-1}).

concentrations of the components are established during the course of the titration.

Acetonitrile is a stronger hydrogen-bond acceptor than dichloromethane, but a weaker acceptor than a pyridine nitrogen (the basic acceptor unit present in **13**), so moderation of the stability of the AAAA–DDDD$^+$ complex is as expected. Dimethylsulfoxide (DMSO) is an even stronger hydrogen-bond acceptor than pyridine and so might be anticipated to eliminate binding entirely. This is the case in pure DMSO, but the AAAA–DDDD$^+$ complex proved remarkably stable in a solution of 10% DMSO in $CHCl_3$ ($K_{12·13} = 3.4 \times 10^5$ M^{-1} ($\Delta G = -31.6$ kJ mol^{-1}); $K_{13·(12·13)} = 1.4 \times 10^5$ M^{-1} ($\Delta G = -29.4$ kJ mol^{-1}); see Fig. 7b and Supplementary Figs S31–S35 for spectra and analysis), despite being held together by only four intercomponent NH···N hydrogen bonds. By way of comparison with other quadruple hydrogen-bonded systems, Meijer and colleagues have reported studies on the tautomerization of **9·9** in $CDCl_3$/DMSO-d_6 mixtures[20]. They determined the complex dimerization constant (K_{dim*}) for **9·9** (equivalent to the product of the K_{dim} of a single tautomer of **9** and the square of the tautomeric equilibrium constant (K_{taut}) and slightly underestimates K_{dim}) to be $\sim 1.7 \times 10^2$ M^{-1} ($\Delta G = -12.8$ kJ mol^{-1}) in 10% DMSO-d_6 in $CDCl_3$ (~ 50 M^{-1} in neat DMSO)[31]. Zimmerman has reported a K_a of 3.0×10^3 M^{-1} ($\Delta G = -19.9$ kJ mol^{-1}) for an ADDA–DAAD complex in 5% DMSO-d_6 in $CDCl_3$ (ref. 22).

The unexpected stability of the AAAA–DDDD$^+$ complex **12·13** in very polar solvent systems such as DMSO/chloroform and, to a lesser extent, acetonitrile, may be a result of poor solvation of the uncomplexed components. In particular, the high density of interaction sites in **12**, in either tautomer, may make it difficult to solvate the uncomplexed DDDD$^+$ partner effectively. Closely packed solvent heteroatoms around the hydrogen-bond donor solvation sites would strongly repel each other. Such ineffective solvation of free hosts or guests is reminiscent of the increase in binding strength found using bulky solvents[32] and in systems with restricted space[33].

Conclusions

The AAAA–DDDD$^+$ quadruple hydrogen-bond complex **12·13** exhibits exceptional complex stability for a small-molecule hydrogen-bond array in a range of solvent systems ($K_a > 3 \times 10^{12}$ M^{-1} in CH_2Cl_2, 1.5×10^6 M^{-1} in CH_3CN and 3.4×10^5 M^{-1} in 10% DMSO/$CHCl_3$). Although the photostability of **13** may be a concern for applications with intense light sources, the strength of binding and ease of synthesis of each partner of the AAAA–DDDD$^+$ hydrogen-bond motif makes it a promising candidate for incorporation into supramolecular polymers and other functional materials.

Methods

Synthesis, characterization and spectroscopic information for all compounds and the details of the UV/vis titration experiments can be found in the Supplementary Information.

Description of UV/vis titration experiments. UV/vis experiments were performed on a Varian Cary Bio-50 UV/vis spectrometer. Glass-distilled solvents were used throughout. Absolute concentrations for each component of a titration experiment were determined by standard calibration curves, and appropriate aliquots added relative to the determined concentrations. All titration experiments with acceptor (A) as the host species were performed with a background concentration of acceptor in the guest solution so as to maintain a constant concentration of host species. Aliquot injections were carried out with Hamilton microsyringes into a quartz cuvette (cuvette size was dependent on concentrations used: 2 mm, 1 cm or 5 cm cells) with a Teflon stopper. Spectrometer settings varied depending on the absorbance and concentrations of the components and can be found for each titration experiment in the Supplementary Information. Spectral analyses were performed using the ReactLab^{TM} Equilibria spectral analyses suite (Jplus Consulting, www.jplusconsulting.com). Repetitions of the binding experiments for each complex gave association constants within 15% of the values shown, except for the analysis of **12·13**, which deviated by 18% (the error in data fitting for each experiment was <5%).

Received 27 August 2010; accepted 13 January 2011; published online 21 February 2011

© 2011 Macmillan Publishers Limited. All rights reserved.

References

1. de Greef, T. F. A. *et al.* Supramolecular polymerization. *Chem. Rev.* **109,** 5697–5754 (2009).
2. Fathalla, M., Lawrence, C. M., Zhang, N., Sessler, J. L. & Jayawickramarajah, J. Base-pairing mediated non-covalent polymers. *Chem. Soc. Rev.* **38,** 1608–1620 (2009).
3. de Greef, T. F. A. & Meijer, E. W. Supramolecular polymers. *Nature* **453,** 171–173 (2008).
4. Wilson, A. J. Non-covalent polymer assembly using arrays of hydrogen-bonds. *Soft Matter* **3,** 409–425 (2007).
5. Zimmerman, S. C. & Corbin, P. S. Heteroaromatic modules for self-assembly using multiple hydrogen bonds. *Struct. Bonding (Berlin)* **96,** 63–94 (2000).
6. Jorgensen, W. L. & Pranata, J. Importance of secondary interactions in triply hydrogen bonded complexes: guanine–cytosine vs uracil-2,6-diaminopyridine. *J. Am. Chem. Soc.* **112,** 2008–2010 (1990).
7. Popelier, P. L. A. & Joubert, L. The elusive atomic rationale for DNA base pair stability. *J. Am. Chem. Soc.* **124,** 8725–8729 (2002).
8. Quinn, J. R., Zimmerman, S. C., Del Bene, J. E. & Shavitt, I. Does the A·T or G·C base-pair possess enhanced stability? Quantifying the effects of CH···O interactions and secondary interactions on base-pair stability using phenomenological analysis and *ab initio* calculations. *J. Am. Chem. Soc.* **129,** 934–941 (2007).
9. Kyogoku, Y., Lord, R. C. & Rich, A. The effect of substituents on the hydrogen bonding of adenine and uracil derivatives. *Proc. Natl Acad. Sci. USA* **57,** 250–257 (1967).
10. Kyogoku, Y., Lord, R. C. & Rich, A. An infrared study of the hydrogen-bonding specificity of hypoxanthine and other nucleic acid derivatives. *Biochim. Biophys. Acta* **179,** 10–17 (1969).
11. Murray, T. J. & Zimmerman, S. C. New triply hydrogen bonded complexes with highly variable stabilities. *J. Am. Chem. Soc.* **114,** 4010–4011 (1992).
12. Bell, D. A. & Anslyn, E. V. Establishing a cationic AAA–DDD hydrogen bonding complex. *Tetrahedron* **51,** 7161–7172 (1995).
13. Djurdjevic, S. *et al.* Extremely strong and readily accessible AAA–DDD triple hydrogen bond complexes. *J. Am. Chem. Soc.* **129,** 476–477 (2007).
14. Blight, B. A. *et al.* AAA–DDD-triple hydrogen bond complexes. *J. Am. Chem. Soc.* **131,** 14116–14122 (2009).
15. Sartorius, J. & Schneider, H.-J. A general scheme based on empirical increments for the prediction of hydrogen-bond associations of nucleobases and of synthetic host–guest complexes. *Chem. Eur. J.* **2,** 1446–1452 (1996).
16. Beijer, F. H., Kooijman, H., Spek, A. L., Sijbesma, R. P. & Meijer, E. W. Self-complementarity achieved through quadruple hydrogen bonding. *Angew. Chem. Int. Ed.* **37,** 75–78 (1998).
17. Sijbesma, R. P. *et al.* Reversible polymers formed from self-complementary monomers using quadruple hydrogen bonding. *Science* **278,** 1601–1604 (1997).
18. Ligthart, G. B. W. L., Ohkawa, H., Sijbesma, R. P. & Meijer, E. W. Complementary quadruple hydrogen bonding in supramolecular copolymers. *J. Am. Chem. Soc.* **127,** 810–811 (2005).
19. Lafitte, V. G. H. *et al.* Quadruply hydrogen bonded cytosine modules for supramolecular applications. *J. Am. Chem. Soc.* **128,** 6544–6545 (2006).
20. Beijer, F. H., Sijbesma, R. P., Kooijman, H., Spek, A. L. & Meijer, E. W. Strong dimerization of ureidopyrimidones via hydrogen bonding. *J. Am. Chem. Soc.* **120,** 6761–6769 (1998).
21. Greco, E. *et al.* Cytosine modules in quadruple hydrogen bonded arrays. *New J. Chem.* **34,** 2634–2642 (1998).
22. Corbin, P. S. & Zimmerman, S. C. Self-association without regard to prototropy. A heterocycle that forms extremely stable quadruply hydrogen-bonded dimers. *J. Am. Chem. Soc.* **120,** 9710–9711 (1998).
23. Park, T., Zimmerman, S. C. & Nakashima, S. A highly stable quadruply hydrogen-bonded heterocomplex useful for supramolecular polymer blends. *J. Am. Chem. Soc.* **127,** 6520–6521 (2005).
24. Kuykendall, D. W., Anderson, C. A. & Zimmerman, S. C. Hydrogen-bonded DeUG·DAN heterocomplex: structure and stability and a scalable synthesis of DeUG with reactive functionality. *Org. Lett.* **11,** 61–64 (2009).
25. Brammer, S., Lüning, U. & Kühl, C. A new quadruply bound heterodimer DDAD·AADA and investigations into the association process. *Eur. J. Org. Chem.* 4054–4062 (2002).
26. Taubitz, J. & Lüning, U. The AAAA·DDDD hydrogen bond dimer. Synthesis of a soluble sulfurane as AAAA domain and generation of a DDDD counterpart. *Aust. J. Chem.* **62,** 1550–1555 (2009).
27. Hisamatsu, Y., Shirai, N., Ikeda, S.-i. & Odashima, K. A new quadruple hydrogen-bonding module based on five-membered heterocyclic urea structure. *Org. Lett.* **12,** 1776–1779 (2010).
28. McGhee, A. M., Kilner, C. & Wilson, A. J. Conformer independent heterodimerisation of linear arrays using three hydrogen bonds. *Chem. Commun.* 344–346 (2008).
29. Connors, K. A. *Binding Constants: The Measurement of Molecular Complex Stability* (Wiley-Interscience, 1987).
30. Hunter, C. A. Quantifying intermolecular interactions: guidelines for the molecular recognition toolbox. *Angew. Chem. Int. Ed.* **43,** 5310–5324 (2004).
31. Lafitte, V. G. H., Aliev, A. E., Hailes, H. C., Bala, K. & Golding, P. Ureidopyrimidones incorporating a functionalizable *p*-aminophenyl electron-donating group at C-6. *J. Org. Chem.* **70,** 2701–2707 (2005).
32. Chapman, K. T. & Still, W. C. A remarkable effect of solvent size on the stability of a molecular complex. *J. Am. Chem. Soc.* **111,** 3075–3077 (1989).
33. Mecozzi, S. & Rebek, J. Jr. The 55% solution: a formula for molecular recognition in the liquid state. *Chem. Eur. J.* **4,** 1016–1022 (1998).

Acknowledgements

The authors thank S. Cockroft (University of Edinburgh) for useful discussions and performing the electrostatic potential computations. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC). D.A.L. and C.A.H. are EPSRC Senior Research Fellows. B.A.B. is a Marie Curie Fellow (IIF) within the European Union 7th Framework Programme.

Author contributions

P.I.T.T. carried out the experimental work. P.I.T.T., B.A.B., D.A.L. and H.M. contributed to the design of the experiments and analysis of the data. C.A.H. designed the complex stability models and calculations. All of the authors contributed to writing the paper.

Additional information

The authors declare no competing financial interests. Supplementary information and chemical compound information accompany this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/. Correspondence and requests for materials should be addressed to D.A.L.

© 2011 Macmillan Publishers Limited. All rights reserved.